Training Manual

Website: www.saspac.org
Email: info@saspac.org
Helpdesk: 020 7983 4348

February 2009

Contents

Introduction **The 2001 Census and accessing datasets through SASPAC…… 3**

Annex1…….……...…….28
Annex2…….31

Module 1 **Selection of variables……..…… 37**

Module 2 **Selection of Tables…..…….……... 49**

Module 3A **Selection of Areas by Text Listing……..…..…….. 57**

Module 3B **Selection of Areas Using Mapping……..…..…….. 63**

Module 4 **Creating New Variables…….……. 75**

Module 5 **Output of Information from SASPAC…...…….…….. 83**

Module 6 **Selection of Areas – Select IF……..91**

Module 7A **Creating New Zones – Individual Area Selection…...…….. 97**

Module 7B **Creating New Zones – Gazetteer Files…….……. 105**

Module 8 **Creating User–Defined Tables……113**

Module 9 **Use of Google™ Maps and SASPAC……..……... 119**

Module 10 **Use of Double Geography datasets within SASPAC (2001 census data)…..……..……... 131**

Appendix 1 **Trainers worked example answers…… 147**

Introduction to the 2001 Census and the use of SASPAC to access the datasets made available by the Census Offices

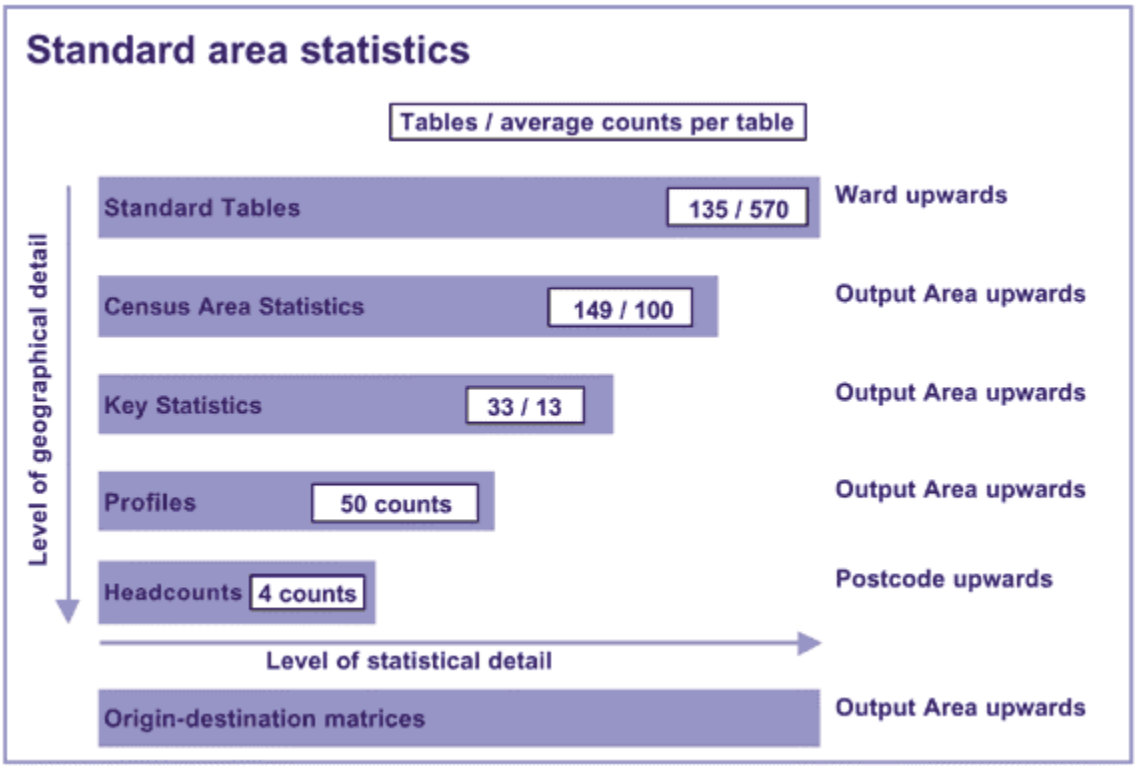
Introduction.

The 2001 Population Census was conducted on 29 April 2001 across the whole of the United Kingdom. The Census is a count of all people and households in the UK. It provides essential statistical information, enabling the planning and funding of public services, including education, health and transport. Results also support research and business. Information is available from the national level down to neighbourhoods.

The first results were made available on 30 September 2002. These consisted of single years of age by sex counts at the national level, with five-year age group data being available at local authority level. This data was followed on 13 February 2003 by the Key Statistics for local authorities, and on May 7 2003 by the Standard Tables for local authorities. During the summer and early autumn of 2003 a continuing stream of data at all area levels and covering a huge number of topics was released by the Census Offices.

The Standard area statistics datasets currently being produced and planned by the Census Offices are as shown below. Note that this is copied from the Census Offices web site at:

http://www.statistics.gov.uk/census2001/op4.asp



'Tables' include 'univariate' counts of single variables as well as cross-tabulations of two or more variables. The average number of counts per table is indicative, but includes totals and sub-totals in tables.

Sets at each level of the hierarchy are sub-sets of those at the levels above, that is the smaller sets nest into the larger. The standard sets in order of increasing detail are

- **Headcounts** - number of people (males/females) and households.

- **Profiles** - a number of standard 'templates', designed as indicators to be presented as percentages in a limited number of simple tables.
- **Key Statistics** - cover all main Census topics, designed as indicators to be presented as percentages in a limited number of simple tables.
- **Census Area Statistics (CS)** - generally presented as cross-tabulations, but also including the simple 'univariate' tables, covering all main Census topics and relationships between them.
- **Standard Tables (ST)** - generally presented in cross-tabulations giving more depth than the CS, including additional 'topic' extension tables for larger populations.
- **Origin/destination** - figures on the people with workplaces in an area and on out-migrants are included in the Standard Tables and CAS, but the flows of workers and migrants between areas are provided in additional matrices.

SASPAC enables access to all these datasets, although the Origin/Destination is a specialist dataset not covered in the basic training modules, while the Headcounts and Profiles are probably better handled through software such as Excel (possibly in combination with Word).

The Output Areas referred to in the above table are small areas created automatically by a computer algorithm from aggregations of unit postcodes (full 7 or 8 character codes), and which have a target population of 120 households in England and Wales, and 40 households in Scotland. There is a more detailed discussion of Output Areas in Module 1 of this Training Manual.

The Census Offices have been planning the content of all these datasets for a number of years, and many revisions were made following consultations with users. However, the contents of the datasets were not completely finalised until the datasets were themselves released to users.

In order to allow easy access to the datasets, SASPAC must contain information about the exact construct and content of each of the datasets in terms of the exact layout of each of the tables to be presented as part of that dataset. As this information was not fully available until the datasets were released, there was a delay between the data becoming available, and the software containing all the information required to recognise and interrogate the 2001 datasets. This will be fully available with the release of version 7 of SASPAC in late November.

This information is added to the software in the form of metadata (termed Framework Files) that are basically descriptions of the textual information that defines each individual table. Extensions of these Framework Files allow the searching facilities to be developed for each dataset.

Knowledge of the full content of each dataset is also required before the data can be loaded into SASPAC as SASPAC System Files.

Apart from the special requirements associated with identification and searching of the 2001 datasets, the functionality of the software applies equally to the 2001 and 1991 data. Thus, all the skills such as file selection, variable selection, and area selection,

along with export and presentation, that can be learnt through use of 1991 data, can equally be applied to the 2001 data.

Basics of SASPAC and the Census

What is SASPAC

SASPAC is the software initially designed specifically to facilitate the storage, interrogation, and analysis of the large quantity of information derived from the 1991 Census of Population, and presented by the Census Offices (OPCS – as it was then known - and GROS) in the form of the Small Area Statistics (SAS) and the Local Base Statistics (LBS). Since its original release, it has been enhanced by the addition of many features, and by porting it from a mainframe environment to a Windows environment on PCs. It has now been developed to ensure compatibility with the various 2001 Census datasets, such as the Standard Tables and the Census Area Statistics, and also to ensure that technological changes do not cause conflict with the software.

What are the Census datasets?

These are datasets derived by the Census Offices from questionnaire responses. They are aggregated data related to geographic areas and not to individual households or persons.

2001 Census – Key Statistics, Census Area Statistics, Standard Tables, Theme Tables, Univariate Tables, Special Workplace Statistics, and Special Migration Statistics.

The 2001 Census data output may be considered as being structured as three distinct datasets, although they will be received from the Census Offices as seven datasets. These are as shown in the following table, which shows the smallest area level for which the data are available, and also the three groups into which they fall.

2001 Census Dataset Name	Smallest Area Availability
Key Statistics	Output Area
Univariate Statistics	Output Area
Standard Tables	Wards
Theme Tables	Wards
Armed Forces	Wards
Census Area Statistics	Output Area
Census Area Theme Tables	Output Area

The 2001 Key Statistics (KS) are a set of tables which are one-dimensional with each table having a set of column headings, with the rows being defined by area. On the other hand, the 2001 Univariate Tables (UV) are also one-dimensional, but with only one column per area, with rows defining variables. As both are one-dimensional they form a group separate from the others.

The 2001 Standard Tables (ST) are a set of cross-tabulations for all local authority wards and higher areas such as local authorities, counties, and regions for the whole of the United Kingdom. They are also available for other areas such as postcode sectors and health areas. In general, the tables are consistent across the whole of the UK, but some tables are only applicable to specific parts of the country.
The ST currently consist of 104 tables containing 60,999 different counts for areas in England and 117 tables containing 65,439 counts for areas in Wales.

The Theme Tables (TT) are a set of tables, each of which is related to a specific theme, such as ethnicity, age, dependent children, or travel to work. Each of these themes is then cross tabulated against a number of separate topics. They are available for the same areas as the ST.

The TT currently consist of 11 tables containing 8,755 different counts for areas in England and 12 tables containing nearly 10,495 counts for areas in Wales.

The 2001 Census Area Statistics (CS) are a set of cross-tabulations for all Output Areas and higher areas such as local authority wards, local authorities, counties, and regions for the whole of the United Kingdom. As with the ST they are also available for other areas such as postcode sectors and health areas. In general, the tables are consistent across the whole of the UK, but some tables are only applicable to specific parts of the country.

The CS currently consist of 71 tables containing 10,448 different counts for areas in England and 75 tables containing 10,761 counts for areas in Wales.

The Census Area Theme Tables (CT) are a set of tables that bear the same relationship to the CS as the TT do to the ST. That is, they each are related to a specific theme, such as ethnicity, age, dependent children, or travel to work. Each of these themes is then cross-tabulated against a number of separate topics.

The CT currently consist of 7 tables containing 2,891 different counts for areas in England and 9 tables containing nearly 4,403 counts for areas in Wales.

Finally, the Armed Forces Tables are a small sub-set of the Standard Tables, that are restricted to armed forces populations.

The Special Workplace and Special Migration Statistics relate to statistics associated with both an origin zone and a destination zone. They are currently under development, and first release of data is expected towards the end of 2003.

1991 Census - SAS, LBS, SMS, and SWS

The 1991 SAS (Small Area Statistics) consist of some 9,000 counts arranged as 83 tables for areas in England, 84 tables for areas in Wales, and 86 tables for areas in Scotland. The SAS dataset is available for Enumeration Districts (ED) in England and Wales, and for Output Areas (OA) in Scotland. It is also available for any higher areas such as wards, local authorities etc.

The 1991 LBS (Local Base Statistics) consist of some 20,000 counts arranged as 95 tables for areas in England, 96 tables for areas in Wales, and 99 tables for areas in Scotland. The LBS dataset is available for wards in England and Wales, and for postcode sectors in Scotland. It is also available for any higher areas such as local authorities, counties etc.

The 1991 SMS (Special Migration Statistics) consist of 90 counts arranged as 11 tables for wards in England and Wales, and for postcode sectors in Scotland. They relate areas of residence at census date to areas of residence one year prior to the census.

The 1991 SWS (Special Workplace Statistics) consist of some 1400 counts arranged as 26 tables for wards in England and Wales, and for postcode sectors in Scotland. The SWS consist of three sets:

Set A: Statistics for areas of residence (9 tables and 574 variables)
Set B Statistics for area of workplace (9 tables and 570 variables)
Set C Statistics linking areas of residence to areas of workplace (8 tables and 274 variables)

What is an Output Area (OA) (or Enumeration District (ED))?

The census questionnaires were distributed and collected in both 1991 and 2001 on the basis of areal units known as Enumeration Districts (EDs). Such areas are defined as the area covered by a single enumerator, and typically would contain 200 households or 450 persons. However, it is possible for EDs to contain very many fewer persons or households. In England and Wales in 1991, the EDs, by default, became the areas for which statistics were output, whereas in Scotland, Output Areas (OA) were created which were not necessarily the same as the enumeration area

For 2001, England and Wales followed the Scottish model, so that OAs are the common base areal unit for output of data across the whole of Great Britain. Each questionnaire returned to the Census Offices has an associated unit postcode (of the form SE1 2AA), and, in theory, data for unit postcodes could be created from responses to the Census questionnaires. However, unit postcodes cover very small areas and would contain too few people to allow the Census Offices to release detailed information without breaching confidentiality restrictions placed on the release of Census data.

An automatic aggregation algorithm was developed at the University of Southampton whereby unit postcodes in England and Wales were merged into successively larger units, until a pre-defined target for the number of households within the area was reached. This target was, in England and Wales, 120 households, although there are some Output Areas with fewer households. There was also a lower threshold for the size of an Output Area, beneath which no data would be released, as described later in the section on restriction/suppression.

In Scotland a similar procedure was followed to create Output Areas, but the target number of households was set much lower at 50 households. Output Areas in Scotland are therefore much smaller than those in England and Wales.

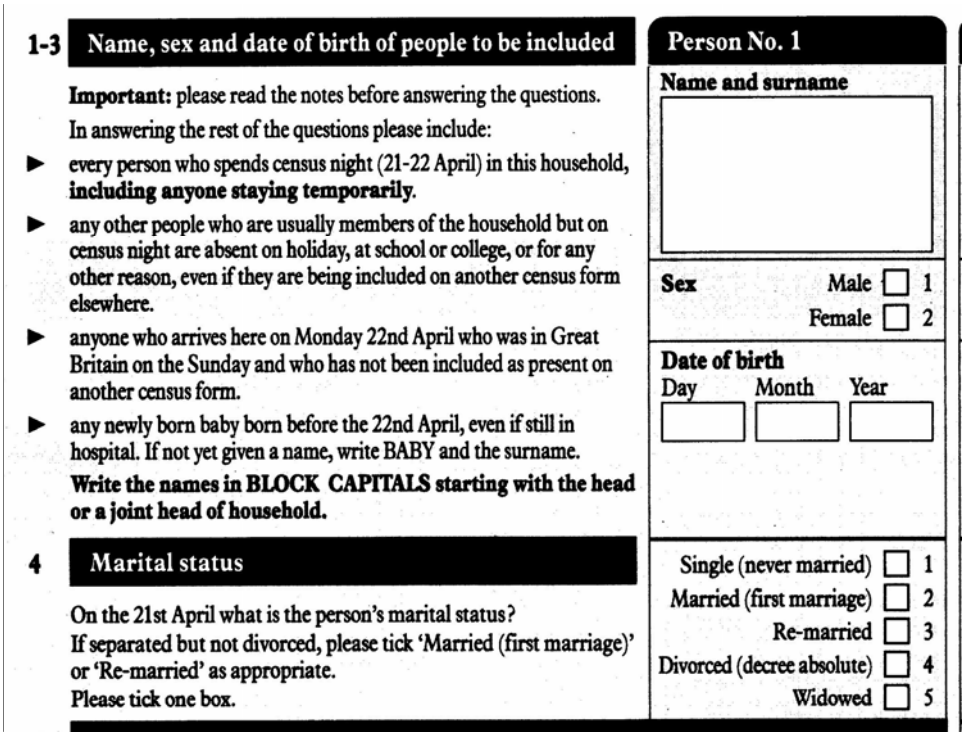
What is a count?

To produce the census statistics, all the information recorded on the census questionnaires - except the names and addresses - is entered onto a database, and responses to the individual questions combined to create multi-variate counts or variables. For example, the number of married males aged 60 to 64 in a given area would be a single count or variable.

Why were there 100% and 10% datasets in 1991.

Some of the counts or variables derived from the census questionnaire in 1991 are considered to be hard to code, and are therefore only presented for 10% of the population. Such a question was number 15, which related to occupation, and this is shown below in Figure 2, in comparison with the questions on sex, date of birth and marital status as shown in Figure 1. During the 100% processing of returns, a sample of 10% of households and persons within them, plus a sample of 1 in 10 persons in

communal establishments, was drawn from the validated records. The data in this sample relating to relationship within families, occupation, industry, workplace, etc was then processed. Tables 71 to 99 (both SAS and LBS) relate to 10% data only.



1-3 Name, sex and date of birth of people to be included

Important: please read the notes before answering the questions.

In answering the rest of the questions please include:

- every person who spends census night (21-22 April) in this household, **including anyone staying temporarily.**
- any other people who are usually members of the household but on census night are absent on holiday, at school or college, or for any other reason, even if they are being included on another census form elsewhere.
- anyone who arrives here on Monday 22nd April who was in Great Britain on the Sunday and who has not been included as present on another census form.
- any newly born baby born before the 22nd April, even if still in hospital. If not yet given a name, write BABY and the surname.

Write the names in BLOCK CAPITALS starting with the head or a joint head of household.

Person No. 1

Name and surname

Sex Male ☐ 1 Female ☐ 2

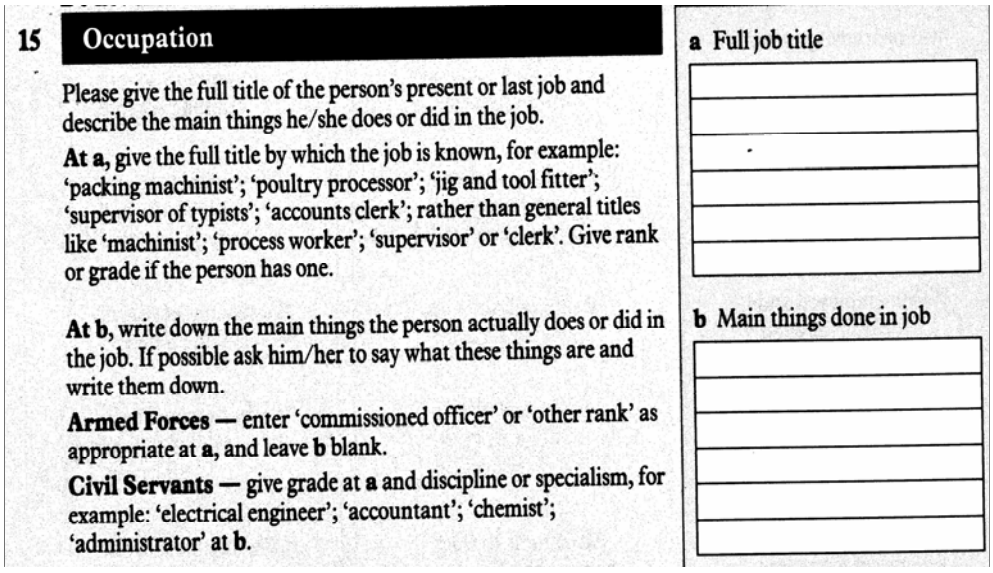
Date of birth Day Month Year

4 Marital status

On the 21st April what is the person's marital status?
If separated but not divorced, please tick 'Married (first marriage)' or 'Re-married' as appropriate.
Please tick one box.

Single (never married) ☐ 1
Married (first marriage) ☐ 2
Re-married ☐ 3
Divorced (decree absolute) ☐ 4
Widowed ☐ 5

Figure 1: 'Easy-to-Code' questions in the 1991 Census questionnaire.



15 Occupation

Please give the full title of the person's present or last job and describe the main things he/she does or did in the job.

At a, give the full title by which the job is known, for example: 'packing machinist'; 'poultry processor'; 'jig and tool fitter'; 'supervisor of typists'; 'accounts clerk'; rather than general titles like 'machinist'; 'process worker'; 'supervisor' or 'clerk'. Give rank or grade if the person has one.

At b, write down the main things the person actually does or did in the job. If possible ask him/her to say what these things are and write them down.

Armed Forces — enter 'commissioned officer' or 'other rank' as appropriate at **a,** and leave **b** blank.

Civil Servants — give grade at **a** and discipline or specialism, for example: 'electrical engineer'; 'accountant'; 'chemist'; 'administrator' at **b.**

a Full job title

b Main things done in job

Figure 2: 'Hard-to-Code' question in the 1991 Census questionnaire.

For 2001, Optical Character Recognition (OCR) techniques were used, which in conjunction with algorithms developed to interpret respondent-generated text, meant that all responses could be coded automatically with manual intervention and validation required in a minority of cases.

What is restriction or suppression?
In order to lessen the risk of inadvertent disclosure of information about identifiable individuals or households in small areas, thresholds are applied to the data, such that the data are not released for areas following below these thresholds.

In 2001, the thresholds applied are 40 households and 100 persons for data at the Output Area level, and 400 households and 1,000 persons for data at the ward level. In Scotland the thresholds at the Output Area level are 20 households and 50 persons, with the ward level thresholds being as in England and Wales.

In 1991, the SAS dataset was not released for areas with fewer than 16 households and 50 persons. Statistics for such areas were merged into contiguous areas provided that the statistics for the amalgamation satisfied the threshold for confidentiality. The corresponding figures for the LBS in 1991 were 320 households and 1000 persons.

What are Special Areas?
In 1991, the Census Offices defined Special Enumeration Districts as communal establishments such as large hotels, hospitals, or defence establishments where 100 or more persons were expected to be present on census night. The characteristics of the populations within SEDs are often markedly different from those in the surrounding area. *Statistics for SEDs are not included in the statistics for the OAs or EDs which include them.*

In 2001, Special Areas do not exist, as all communal establishments are considered as part of the standard Output Area in which they are located.

What are Shipping Areas?
In 1991, for Census purposes, each local government district included a shipping ward (or postcode sector in Scotland). This shipping area was created for the purposes of enumerating persons on board ships (excluding houseboats) in transit. For each shipping ward or postcode sector, there was at least one shipping ED or OA.

In 2001, Shipping Areas do not exist, as all such craft are considered as part of the standard Output Area in which they are located.

How does SASPAC operate?
The operation of SASPAC for Windows is controlled by a text file - SASPAC.INI. In general, users will not need to amend this, but would be advised to be aware of it. It may require amendment at installation of the software, but should not need further changes. This file defines defaults for the working parameters of SASPAC, and specifies default directories for the files necessary for the operation of SASPAC. A copy of an example initialisation file is shown in Annex 1 to this introduction.

Most of the parameters specified in this file can be amended through SASPAC by use of the Tools/SASPAC Configuration menu. Use of this option results in a dialog box as shown in Figure 3.

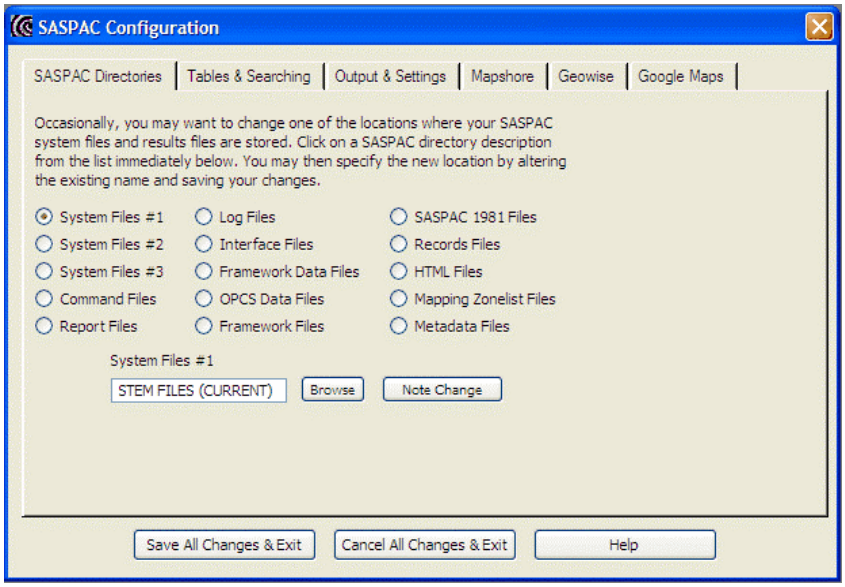


Figure 3: SASPAC Configuration dialog box

Initially, SASPAC reads the raw SAS and LBS as distributed by the Census Offices, and converts and compresses the data into an internal format which is more efficient in terms of data storage, and enables easy access to the statistics. The internal format files containing the compressed, converted data are referred to as SASPAC System Files. Once the conversion has been undertaken, users have no further need for the loading capabilities of SASPAC. The 2001 Census datasets are distributed to users as pre-loaded System Files.

What are SASPAC System Files?

The System files may be considered to be very large matrices or tables where the areas held on the file identify the rows of the table, and the variables available for each of those areas identify the columns of the table. This is represented schematically in Figure 4 below.

	Table 1			Table 2			.....	Table m		
	var 01	.....	var a	var 01	.....	var b	.....	var 01	.....	var c
Zone 1							.....			
Zone 2							.....			
Zone 3							.....			
Zone 4							.....			
Zone 5							.....			
Zone 6							.....			
.							.....			
.							.....			
.							.....			
.							.....			
.							.....			
.							.....			
.							.....			
.							.....			
.							.....			
Zone n							.....			

Figure 4: Schematic representation of SASPAC System File.

The Census Offices, in defining the contents of the various datasets, identify the variables available for each area as a set of tables.

How are Tables identified within SASPAC?

When the Census data is loaded into SASPAC, unique identifiers are allocated to all the tables and variables within the datasets received from the Census Offices. These identifiers then comprise part of the metadata which facilitates the retrieval of individual tables, variables, and areas.

Both the 1991 and 2001 datasets consist of a number of tables, each table containing a number of variables. There are slight differences in the numbers of tables available for areas in England, Wales, and Scotland. The tables below provide summaries for 1991 and 2001 for the major datasets.

2001 ***(estimated)***

	Census Area Statistics			Standard Tables		
	England	Wales	Scotland	England	Wales	Scotland
100% only	76	78	83	112	117	116

1991

	SAS			LBS		
	England	Wales	Scotland	England	Wales	Scotland
100%	61	62	65	66	67	70
10%	21	21	21	29	29	29
Total	82	83	86	95	96	99

In order to allow access to individual tables, as defined by the Census Offices, each of the tables must have unique identifiers. SASPAC does this by allocating a different alphanumeric code, to each table. This alphanumeric code is created in the following way:

2001 Datasets

1st two characters - valid codes KS, UV, CT, CS, ST, TT, AF.

- KS indicates table is from the Key Statistics.
- UV indicates table is from the Univariate Tables.
- CT indicates table is from the Census Area Statistics Theme Tables.
- CS indicates table is from the Census Area Statistics.
- ST indicates table is from the Standard Tables.
- TT indicates table is from the Standard Tables Theme Tables.
- AF indicates table is from the Armed Forces Tables.

3rd to 5th characters - valid codes 001 to 999.

- 001 indicates table 1 within dataset identified by first two characters.
- 002 indicates table 2 within dataset identified by first two characters.
- etc. to
- 999 indicates table 999 within dataset identified by first two characters.

In general, there are no national variations, whereby a table can have different contents depending on the part of the UK to which they refer. Where a similar table is produced for areas in different nations, different table numbers apply. Generally, tables with numbers in the range 1 to 100 are applicable to the whole of the UK, those in the range 101 to 200 apply to areas in England and Wales, those in the range 201 to 300 apply to areas in Scotland and those in the range 301 upwards apply to Northern Ireland. However, at a late stage in the production process, GROS changed this system, and amended tables in the range 1 to 68 so that there are now Scottish variants. This means that within SASPAC, such tables will need to be identified separately, and this is done by using the ‘hundreds’ location of the table number for the ‘S’ identifier. Thus while

there is a table CS022 in England and Wales, the equivalent table in Scotland is identified as table CSS22.

A full set of valid table identifiers is shown in Annex 2 to this introduction.

Note that in 2001 the table identifier must be given five characters. Therefore, where a table lies in the range 1 to 9, two leading zeros have to be introduced. SASPAC will not recognise the code ST5, but will recognise ST005 (or maybe STS05).

1991 Datasets

1st character - valid codes L or S.

- L indicates table is from the LBS.
- S indicates table is from the SAS.

2nd and 3rd characters - valid codes 01 to 99.

- 01 indicates table 1 within dataset identified by first character.
- 02 indicates table 2 within dataset identified by first character.
- etc. to
- 99 indicates table 99 within dataset identified by first character.

In some instances, tables exist only for areas in Wales or Scotland, or for Great Britain as a whole. In other cases, tables for Scotland contain minor differences from those for the rest of GB. In these instances, a fourth character is introduced to the Table identifier, to indicate a national variation. This extra character - S, W, or G - is placed after the dataset identifier, but before the table number.

Thus table 15 of the LBS whose format is different for areas in Scotland from that for areas in England and Wales would be identified as follows:

- L15 Table 15 of the LBS for an area in England or Wales.
- LS15 Table 15 for an area in Scotland.

Note that in 1991 the table identifier must be given three (or four) characters. Therefore, where a table lies in the range 1 to 9, a leading zero has to be introduced. SASPAC will not recognise the code L5, but will recognise L05.

As each table is discretely identified within the dataset, and therefore within the SASPAC System Files, individual tables may be specified for access within SASPAC as portrayed in Figure 5 below.

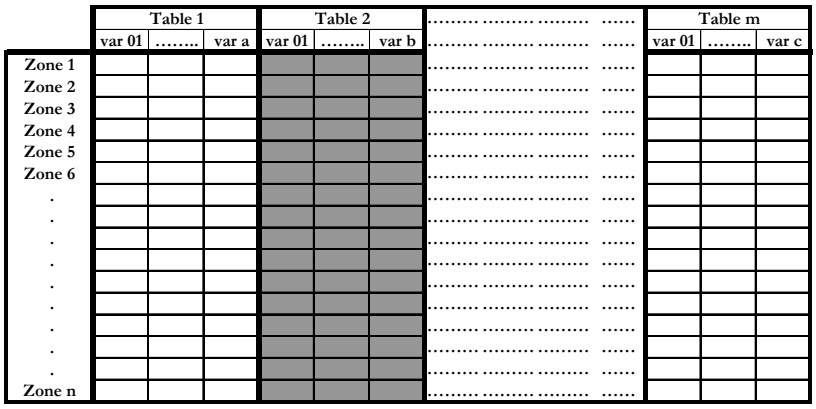


Figure 5: Extraction of individual table(s)

How are Variables identified within SASPAC?

As already mentioned, the SAS and LBS datasets are both defined in terms of a number of tables, each table containing a number of variables. A corollary of the variation in numbers of tables already referred to is that there are slight differences in the numbers of variables available for areas in England, Wales, and Scotland. The table below provides a summary:

2001 **(estimated)**

	Census Area Statistics			Standard Tables		
	England	Wales	Scotland	England	Wales	Scotland
100% only	11,000	11,500	12,000	61,000	66,000	68,000

1991

	SAS			LBS		
	England	Wales	Scotland	England	Wales	Scotland
100%	6,501	6,837	7,442	15,522	15,858	17,019
10%	2,498	2,498	2,538	4,807	4,807	4,889
Total	8,999	9,335	9,980	20,329	20,665	21,908

In order to allow access to individual variables, as defined by the Census Offices, each of the variables must have a unique identifier. SASPAC does this by allocating a different alphanumeric code which is an extension of the table code, to each variable. This alphanumeric code is created in the following way:

2001 Datasets

1st two characters - valid codes KS, UV, CT, CS, ST, AF, TT.

- KS indicates table is from the Key Statistics.
- UV indicates table is from the Univariate Tables.
- CT indicates table is from the Census Area Statistics Theme Tables.
- CS indicates table is from the Census Area Statistics.
- ST indicates table is from the Standard Tables.
- AF indicates table is from the Armed Forces Tables.
- TT indicates table is from the Standard Tables Theme Tables.

3rd to 5th characters - valid codes 001 to 999.

- 001 indicates table 1 within dataset identified by first two characters.

002 indicates table 2 within dataset identified by first two characters.
etc. to
999 indicates table 999 within dataset identified by first two characters.

6th to 9th characters - valid codes 0001 to 9999.
0001 indicates variable 1 within table identified by first five characters.
0002 indicates variable 2 within table identified by first five characters.
etc. to
9999 indicates variable 9999 within table identified by first five characters.

The note on Scottish variations on Table identifiers also applies to Variable identifiers, as the Table code is contained within the Variable code.

1991 Datasets
1st character - valid codes L or S.
L indicates variable is from the LBS.
S indicates variable is from the SAS.

2nd and 3rd characters - valid codes 01 to 99.
01 indicates table 1 within dataset identified by first character.
02 indicates table 2 within dataset identified by first character.
etc. to
99 indicates table 99 within dataset identified by first character.
4th to 7th characters - valid codes 0001 to 9999.
0001 indicates variable 1 within table identified by first three characters.
0002 indicates variable 2 within table identified by first three characters.
etc. to
9999 indicates variable 9999 within table identified by first three characters.

Note that in both 1991 and 2001, the variable identifier must be given four characters. Therefore, where a variable lies in the range 1 to 999, leading zeroes have to be introduced. For 2001 data, SASPAC will not recognise the code ST00557, but will recognise ST0050057.

Individual variables may be specified for access within SASPAC as portrayed in Figure 6 below.

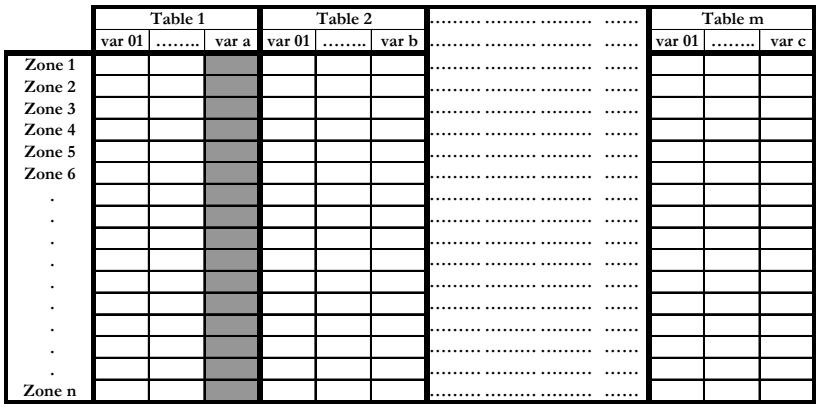


Figure 6: Extraction of individual variable(s)

How are Areas identified within SASPAC?

The manner in which the standard areas for which census information is available are identified is determined by the Census Offices. Each of these areas is allocated a unique identity code (referenced by SASPAC as ZONEID) within a hierarchical structure. The basic areal unit for release of statistics from the 2001 Census in England and Wales is the Output Area (OA). Each OA has been allocated a 10-character reference from which the encompassing ward, district and county (where applicable) identity codes may be deduced. For example, the 2001 OA identified by the code 46RSFA0001 is the first OA within ward 46RSFA, which itself lies within district 46RS, which in turn lies within county 46.

The area coding convention in 2001 is very similar to that used in 1991, but with a very important difference in Scotland, and with a discontinuation of the hierarchy in England and Wales because of the existence of Unitary Authorities.

In Scotland, the Census Office took the decision that Output Areas should nest within local authority districts only, and not within wards (or postcode sectors), while in England and Wales, ward boundaries were recognised when the Output Areas were defined. This means that Output Area codes within Scotland do not contain a code relating to a higher area other than the local authority district, while those in England and Wales maintain the hierarchy as used in 1991.

However, following the creation of Unitary Authorities in England and Wales, and the abolition of some counties, the standard hierarchy is now no longer available. For example, in 1991, Portsmouth was part of the county of Hampshire in a two-tier local authority structure, but by 2001 it had become a Unitary Authority with a single tier structure and was not part of the county. Areas within Portsmouth now do not have the Hampshire code as part of their codes. In 2001 all areas within all Unitary Authorities have a county or Higher level code of 00. Thus any area within Portsmouth will have a code starting with 00, as will areas in Middlesbrough which is also a Unitary Authority. Additionally, areas within the former Metropolitan Counties, e.g. Greater London, or Tyne and Wear, will also obey the same rule.

Output Areas are identified by 10 characters which are allocated as follows:

Characters	Area Identified	Values
1 – 2	County or UA indicator	00 for UAs or 09 to 46
3 – 4	Local Authority	Alphabetic
5 – 6	Ward	Alphabetic
7 – 10	Output Area	0001 to 9999

Other areas such as regions in England, Metropolitan Counties and 1991 Counties are identified by pre-defined codes

1991 Area Coding.
Across England and Wales, the coding structure is as follows:

Counties Coded in the range 01 to 55
Districts Coded in the range AA to TT
Wards Coded FA, FB, FC, FD, etc. within district
EDs Coded 01, 02, 03, 04, etc. within ward.

It should be noted that as districts are uniquely identified across the whole of England and Wales, the county code is not strictly required to identify a district. However, wards are only uniquely identified within districts, and so the district code is necessary to uniquely identify a ward. Similarly for EDs, the district code is required. The rule for using the ZONEID within SASPAC is therefore:

County level	2 character code - e.g. 46
District level	2 character or 4 character code - e.g. 46RS or RS
Ward level	4 character or 6 character code - e.g. 46RSFA or RSFA
ED level	6 character or 8 character code - e.g. 46RSFA01 or RSFA01

Within Scotland there is a small variation to this convention which allows for comparisons with 1981 Census areas. Firstly, the basic areal units are referred to as Output Areas (OAs) rather than Enumeration Districts, and these have been designed to aggregate to 1981 Census EDs. In order to enable this aggregation, Scottish OAs have a ninth (or seventh character) in the ZONEID. This extra character is an alphabetic character which follows the ED identifier.

Thus a valid code for a Scottish OA in 1991 would be 5601AB01A. In this example, the OA identified by the code 5601AB01A is an OA within postcode sector 5601AB, which itself lies within district 5601, which in turn lies within county 56.

Across Scotland, the coding structure is as follows:

Regions	Coded in the range 56 to 67
Districts	Coded in the range 01 to 56
Postcode sectors	Coded AA, AB, AC, AD, etc. within district
1981 EDs	Coded 01, 02, 03, 04, etc. within postcode sectors.

As in England and Wales, districts are uniquely identified across the whole of Scotland. The region code is therefore not strictly required to identify a district.

Note that whereas the region code within Scotland follows on from the end of the range in England and Wales, the district code is a numeric rather than an alphabetic one, and starts from 01. However, as both the region and district code are both numeric, it is usually advisable to use the full code. In addition, the ward level zone is in fact a postcode sector, and codes do not follow the ranges used in England and Wales. Within Scotland, the rule for using the ZONEID within SASPAC is thus:

Region level	2 character code - e.g. 63
District level	2 character or 4 character code - e.g. 6343 or 43
Postcode sector level	4 character or 6 character code - e.g. 6343AE or 43AE
1981 ED level	6 character or 8 character code - e.g. 6343AE14 or 43AE14
OA level	7 character or 9 character code - e.g. 6343AE14A or 43AE14A

In order to identify areas for interrogation and analysis, users will need to be accustomed to the hierarchical structure of the area coding convention.

Using the area code or identifier, it is possible for SASPAC to extract individual areas as depicted in Figure 7 below.

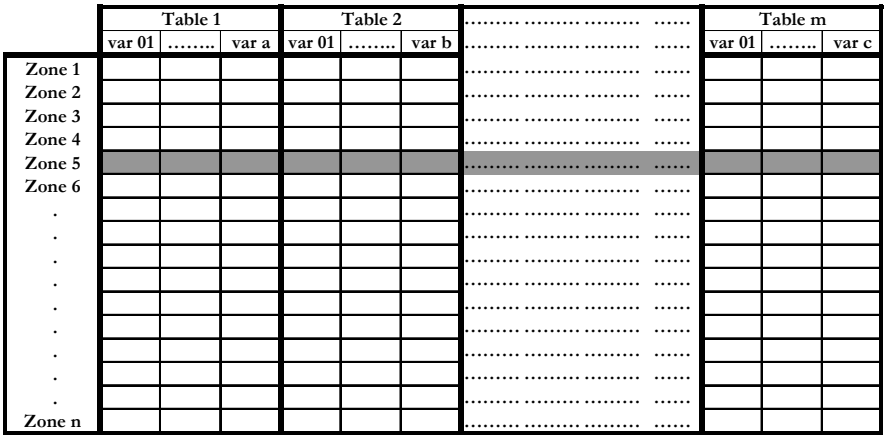


Figure 7: Extraction of individual area(s)

The techniques of selecting individual areas and variables / tables may be combined as indicated in Figures 8 and 9 below. In this way, SASPAC enables users to extract single items of information from the mass of statistics contained within a single System File. Note that within a single task, it is not possible to select one set of variables for one area, and a second set of variables for another.

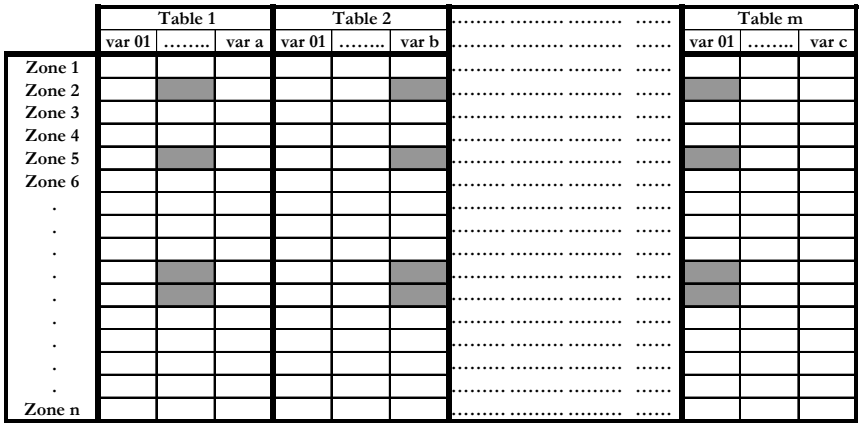


Figure 8: Extraction of individual tables for selected area(s)

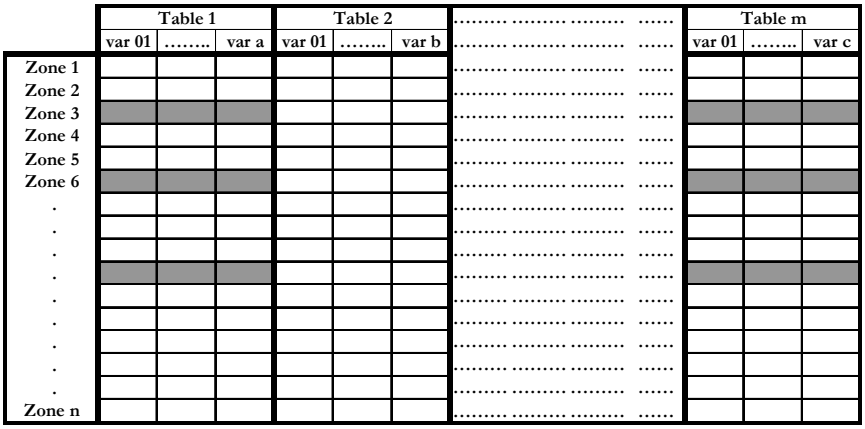


Figure 9: Extraction of individual variables for selected area(s)

How can the contents of a System File be deduced?
Since these System Files are held in a format which is internal to SASPAC, (that is they can only be read by SASPAC), their contents cannot be examined by using a text editor or viewer. There will be occasions when a user is uncertain as to the contents of a system file, and may need some means of examining its contents prior to running a SASPAC task.

Within SASPAC for Windows, there is a facility for doing this, which is accessed through the Tools / System File Details option as shown in Figure 10.

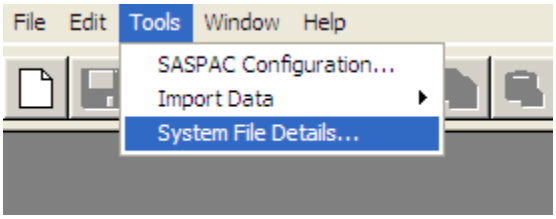


Figure 10: Tools / System File Details selection.

Selection of this option will lead to the user being presented with a dialog box as detailed in Figure 11.

Browse Button used for scanning system files in other directories or drives.

Option buttons for selecting level of detail required for the selected System File.

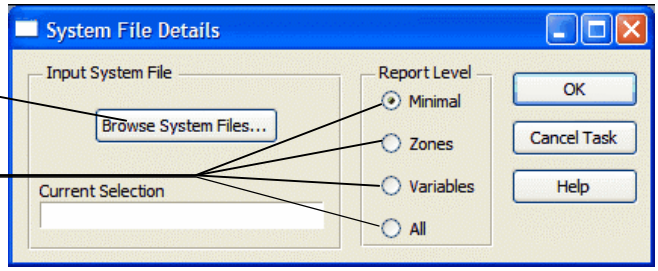


Figure 11. System File Details Dialog Box.

Is there a convention for naming System Files?
Users of SASPAC in a 16-bit environment - Windows 3.1 for example, were constrained by the naming convention of that system whereby each file name has two components - the root, and the extension. The root could have up to 8 characters taken from the standard alphabetic or numeric character sets, while the extension could have up to 3 characters taken from the same sets. Current releases of the software do not operate in a 16-bit environment, and are only operational under Windows 2000 or later, including, XP and Vista.

In a 32-bit environment, there are no such constraints. A file name may have up to 255 characters including embedded blanks so that fully descriptive names may be applied to files. It is suggested that each file name still makes use of the extension as used in a DOS environment to allow easy identification of type of file, not only by the user, but also by the system. The extension should indicate the type of file it refers to, and SASPAC retains a default set of extensions which apply to the different types of files which it accesses. In general, users will not be required to allocate extensions to files, since SASPAC will do this depending upon the context in which the file is created.

Most users however, will still be using some files created under a 16-bit environment, and therefore will be using files with names such as L9123WDH.SYS. The root is normally allocated by the user, and was usually created in such a way that it presents

information as to the content of the file. By allocating meaningful file names, subsequent access and housekeeping would be made easier.

A suggested coding frame for file names in 1991 took account of the following:

To identify:	Suggested code:
Source of data (LBS, SAS, or other)	L, S or as necessary.
Year to which data refers	91, 81, or as necessary.
Area covered by dataset	County code to indicate county, district code to identify district or as necessary.
Area level of dataset	C to indicate county, D to indicate district, W to indicate ward, etc.
Sample level of data	H or T to indicate 100% or 10%

Thus, a suitable root name for a file containing 100% 1991 Census SAS data for all EDs in the (1991) county of Hereford and Worcester (which had a county code of 26) would be:

S9126EDH

If a standard convention for naming files, especially system files, is maintained within an organisation, then all users within that organisation should have the means of identifying the contents of a file from its name.

2001 Census System Files to be supplied centrally by SASPAC will use long file names, and should contain enough information to allow users to select their appropriate files.

What other type of file does SASPAC use?

In order to function efficiently, SASPAC needs to have access to several different types of file. Each of these file types is identified by a default extension to the name. Users are strongly advised not to change these extensions, since SASPAC will not recognise them if they have non-standard extensions.

File type	**Default extension**	**Purpose, content, or use of file.**
Command files	**.CMD**	These files are crucial to all SASPAC operations. They are text files which contain a series of commands or instructions for SASPAC to implement. Each command occupies a single line or a series of linked lines. **Each line must not contain more than 80 characters.** They may be created within SASPAC using the menu system or editor, or they may be created externally via a text editor or word processor.
Log files	**.LOG**	Each time SASPAC acts on a command file, a .LOG file is produced. This file contains information concerning the implementation of the command files, and in the case of

		problems, will report warning and error messages to alert the user.
Report files	**.PRN**	If the objective of a command run is to produce printed output, that output will be written to a .PRN file which may be immediately printed or stored for later output.
Interface files	**(various)**	If the objective of a command run is to produce output for further analysis by other software, or for integration in a published report, such output will be written to an interface file. The particular type of interface file used will depend on the software to which the data is to be exported.
Framework Data files	**.FWD**	These are ASCII representations of the table layouts necessary to produce the Census Offices table images.
Framework files	**.FWK**	Internal format versions of the .FWD files. They supply the text necessary to produce tables, and define the location of data within those tables. The relationship between .FWD and .FWK files is similar to that between .TXT and .SYS files.
Configuration file	**.CFG**	**(No longer used)** A file which contained default parameters enabling SASPAC to operate and to 'know' the location of the types of files it had to access. This file is now incorporated within the initialisation file.
Initialisation file	**.INI**	Windows initialisation file. This file may be modified and customised by users.
Gazetteer files	**.GAZ**	These files contain definitions of new areas or zones in terms of existing areas or zones, and are used in the creation of datasets for new areas as defined by users.
Data Definition Language file	**DDL**	These files define the content and format of the .TXT files which are to be converted to .SYS files, and are only used at the time of creation of the .SYS files.
OPCS Text files	**.TXT**	Raw data files as supplied by the Census Offices, or from other sources.
Transfer files	**.TRF**	Used for transfer of complete (or partial) datasets between SASPAC on two different hardware platforms.
SASPAC4 files	**.S81**	These files are the 1981 SASPAC equivalents of the .SYS files. They may be converted to .SYS files for use with SASPAC(91).

SASPAC will know the content of each of these file types from the extension used, and the default location of each will be given in the initialisation file. Thus, the user will generally only need to know and supply the root file name, since SASPAC will supply the path and extension.

As with most Windows applications, the SASPAC initialisation file - SASPAC.INI – is a file that is accessed by the system every time the application is opened.

Paths defined within the SASPAC initialisation file may be changed either through the appropriate window (as described on page 7), or through use of an external text editor, such as Notepad. However, it is strongly advised that no changes are made within a multi-user system without the prior agreement of all users.

What is a Command File?

The Command File is at the core of all SASPAC operations. Command Files contain sequences of instructions which enable the software to carry out the user's requirements, and are needed for all SASPAC tasks.

Command files may be created interactively through the various window menus described in the SASPAC for Windows User Manual, in which case the user needs to know little, if anything, about the syntax of the command file. Alternatively, they may be created / amended using a text editor, either within SASPAC, or externally. In this case the user will be required to have some knowledge of the command language syntax.

As already stated, a command file is simply a list of instructions for the software to undertake. These instructions are entered as a series of lines, none of which may be greater than 80 characters. The user need not be concerned with this artificial limitation, as SASPAC will handle appropriate breaks and use a '/' character to signify continuation across two or more physical lines. There is a pre-defined logic to the order in which these instructions are given to the software - some are mandatory, while others are optional. Command files created through window menus will automatically be in the correct order *although the options need not have been selected in the correct order*.

Every SASPAC run or command file must contain INPUT and OUTPUT, but between these operations, there are many options which users may take. The standard structure of a command file is:

INPUT	*Mandatory*
Selection	*Optional*
Manipulation	*Optional*
OUTPUT	*Mandatory*

For the purpose of letting SASPAC know that a particular task has been completed, the command END is used. This allows for several tasks to be concatenated so that the structure of the command file becomes:

INPUT	*Mandatory*
Selection	*Optional*
Manipulation	*Optional*
OUTPUT	*Mandatory*
END	*Mandatory*
INPUT	*Mandatory*
Selection	*Optional*
Manipulation	*Optional*
OUTPUT	*Mandatory*
END	*Mandatory*

A final command is mandatory to instruct SASPAC that the command file is completed, and that is the FINISH command. The complete structure of a command file thus becomes:

INPUT	*Mandatory*
Selection	*Optional*
Manipulation	*Optional*
OUTPUT	*Mandatory*
END	*Mandatory*
INPUT	*Mandatory*
Selection	*Optional*
Manipulation	*Optional*
OUTPUT	*Mandatory*
END	*Mandatory*
FINISH	*Mandatory*

Note that the command END may only be followed by an INPUT command or by the FINISH command, (or by the non-executable comment command). In creation of command files through window menus, the END and FINISH commands are automatically added by the system.

The window menus system allows for the creation of single task command files. Multi-task command files may be created by creating the required number of single task command files, and using the Editor to combine these into a single command file, *ensuring that there is only one* FINISH *command.*

Comments within SASPAC command files are denoted by the asterisk character '*' placed in the first position of the command line. It is good practice to place comments within every command file, as this will make it easier to understand their purpose when used by others or when examined after a period of time.

A simple command file for printing a single variable would have the following format:

```
INPUT SYSTEM FILE = L9146WDH
PRINT VARIABLES L020001
OUTPUT PRINT FILE = VARS_L02
END
FINISH
```

This command file instructs SASPAC to read an input file (L9146WDH), to access variable 1 in table 2 of the LBS, and to print this to an output file (VARS_L02). The final two lines are necessary to indicate the end of the task, and the end of a set of tasks.

SASPAC expects to see the command lines written in pre-defined formats or syntax. When creating command files through the window menus system, users need not be aware of these formats, but if existing files are to be edited, or if files are to be created through a text editor, then knowledge of the command line formats is essential.

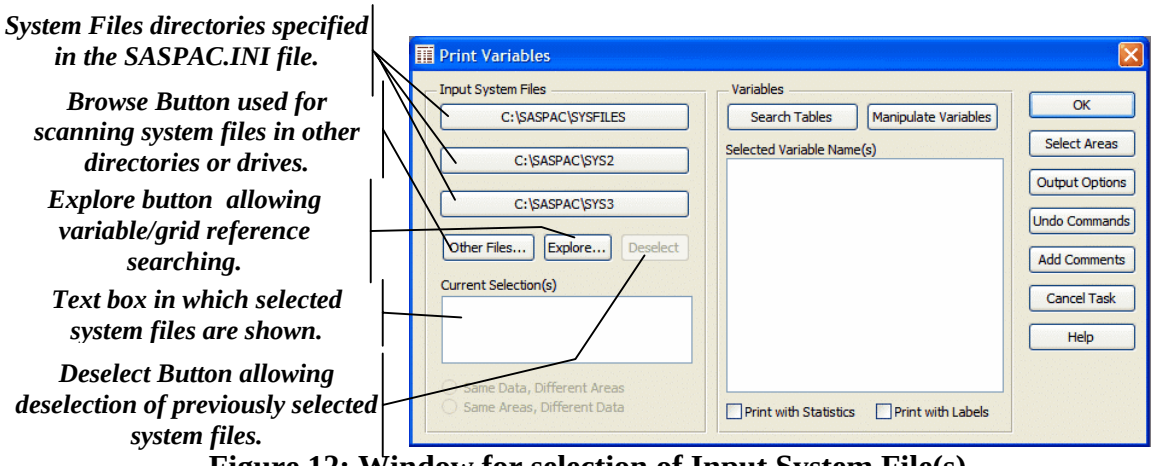


Figure 12: Window for selection of Input System File(s).

The first executable element of any Command File must be an Input System File, and the choice of the input system file is offered to the user through a window similar to that shown in Figure 12.

The exact content of this window will depend on the type of task (and the resultant output) that is to be undertaken.

Why does SASPAC use Frame work Files?

SASPAC System Files contain only data so that if there is a need to output tables in the image of the Census Offices layouts or in a format as defined by users, there is a necessity to merge the data with the appropriate text which defines the tables. This is done through use of Framework Files. These Framework Files are held by SASPAC in a format that is internal to SASPAC, i.e. they cannot be interrogated or edited by other software. In order that users may edit these files, or create their own, ASCII files called Framework Data Files are used. These may be edited, and then loaded into SASPAC for saving as Framework Files. The relationship between Framework Data Files and Framework Files is the same as that between Raw Data Files and SASPAC System files.

The main Framework Files distributed with SASPAC for accessing the 1991 Census datasets were:

TLBS132	To allow printing LBS tables at a maximum of 132 characters per line.
TLBS160	To allow printing LBS tables at a maximum of 160 characters per line.
TSAS132	To allow printing SAS tables at a maximum of 132 characters per line.
TSAS160	To allow printing SAS tables at a maximum of 160 characters per line.
TSWS	To allow printing SWS tables (Sets A & B) at a maximum of 160 characters per line.
SWSC91	To allow printing SWS tables (Set C) at a maximum of 160 characters per line.
SMS91	To allow printing SMS tables at a maximum of 160 characters per line.

Seven Framework Files are currently available for the 2001 Census datasets – one each for the ST, TT, CS, CT, and UV, and two for\the KS. The reason that there are two

Framework Files for the Key Statistics is that as originally designed by the Census Offices, the KS did not fit into the 'standard' for table layouts. Thus there is a framework which adapts the KS as a 'standard' layout with a single area being output to a page, and there is a new variation which allows for areas to appear as the row variable in a table. These areas are user selectable, and this framework allows the user to replicate the layouts for the KS as they appear in published volumes.

How do I input multiple system files?

On most occasions, the areas or data required by a user will be found on a single system file. However, there will be instances when the input of more than one system file is demanded. This might occur, for example, when enumeration districts in more than one county are being compared, or when data is required from both the 100% and 10% datasets in 1991.

SASPAC allows up to nine separate System Files to be identified for inclusion in a single task. If more than one file is to be accessed, SASPAC needs to know how the files are to read in relation to each other.

In combining files within the same task, the following options are allowed:

- same areas, different data; or
- same data, different areas.

Note that the *'Different areas, different data'* combination is not permitted.

In defining 'same' in this context, the rule is that the sub-set of areas or data selected for manipulation must appear on all files. If no sub-set is selected, then the areas or data found on the first file, must appear on all following files.

As already stated, system files may be considered as matrices, where the rows are the areas, and the columns the data. In simplistic terms, SASPAC reads these files by progressing along a row to the end, before proceeding to the next row and repeating the process through to the end of the file.

If more than one file is input, SASPAC needs to be told whether the areas or the variables are the common factor, in order that it may place the appropriate information as an extension of a row (i.e. the areas are common), or as an extension of a column (i.e. the variables are common).

This is done through the pair of complementary commands:

READ IN SERIES
READ IN PARALLEL

If more than one system file is input, the READ IN command must be present in one form or other, and must immediately follow the last INPUT SYSTEM FILE command (except for the possible inclusion of comments).

Files within a single task may be read either in series or in parallel. The two procedures may not be mixed.

i) READ IN SERIES.

If the files to be input contain the same data for different areas, then effectively, SASPAC must add each file onto the end of the previous one to create a matrix which has the same number of columns as the original, but a number of rows equal to the combined total of rows in all files read. This procedure is termed 'Reading in Series' since SASPAC is placing files in a vertical series.

An example of the use of the 'READ IN SERIES' command would be the input of the CS for OAs in a county, along with the CS for OAs in a different county.

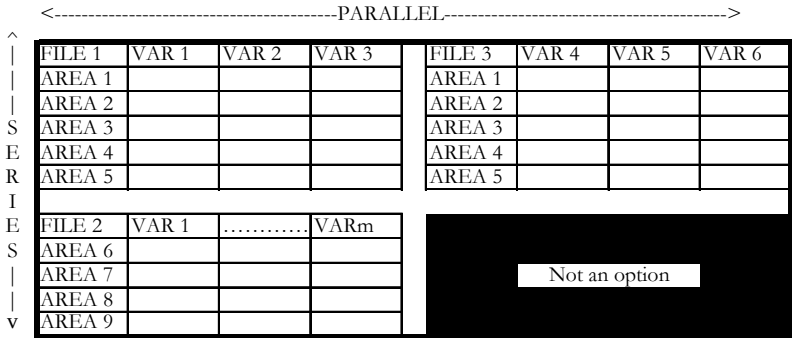
ii) READ IN PARALLEL.

If the files to be input contain different data for the same areas, then effectively, SASPAC must place each file alongside the previous one to create a matrix which has the same number of rows as the original, but a number of columns equal to the combined total of columns in all files read. This procedure is termed 'Reading in Parallel' since SASPAC is placing files alongside (or parallel to) the previous one.

An example of the use of the 'READ IN PARALLEL' command would be the input of the 1991 100% SAS for EDs in one county, along with the 1991 10% SAS for EDs in the same county.

Figure 13 below shows how the two methods of reading System Files operate. If the files to be read contain the same data for different areas, then the two (or more) files are effectively put one on top of the other, and read in series. Conversely, if the files to be read contain different data for the same areas, then they are placed one alongside the other, and read in parallel.

SASPAC for Windows allows the appropriate selection to be made, without the user having to consider whether the required keyword is Series or Parallel. All that is required of the user is the need to identify whether it is the areas or the variables that are common.



FILE 1	VAR 1	VAR 2	VAR 3	FILE 3	VAR 4	VAR 5	VAR 6
AREA 1				AREA 1			
AREA 2				AREA 2			
AREA 3				AREA 3			
AREA 4				AREA 4			
AREA 5				AREA 5			
FILE 2	VAR 1	.....	VARm				
AREA 6							
AREA 7							
AREA 8							
AREA 9							

SERIES

Not an option

Figure 13: Input of more than one System File.

How are tasks created through Windows?

Each of the modules in this training document are related to the optional facilities mentioned on pages 13 and 14. That is, they relate to the selection and manipulation of the data. With the exception of module 7 - Creating New Zones - each module may lead to any of the types of output available within SASPAC. New Zone creation must lead to the derivation and output of a new system file.

In order for SASPAC to perform a task, it must be provided with certain information by the user. To do this the user must ask and answer questions such as:

What type of output is required and where is it to be written to? The answer to this question will define the type of task that is to be created, and will need to be the first consideration.

What data is required - is it SAS or LBS? The answer to this may depend on the answer to the next question?

What area level is required - is it ED / OA, ward / postcode, district, or county / region?

What variables (or tables) are required - do they exist or will they need to be created by the user?

What areas are to be analysed - do they exist or will they need to be created by the user?

How are the areas to be selected - by name, by geography, or by condition?

Once the answers to these questions are available, the user may start the creation of the command file for the task by acting on the response to the first question and using the File / New Task menu option as shown in Figure 14.

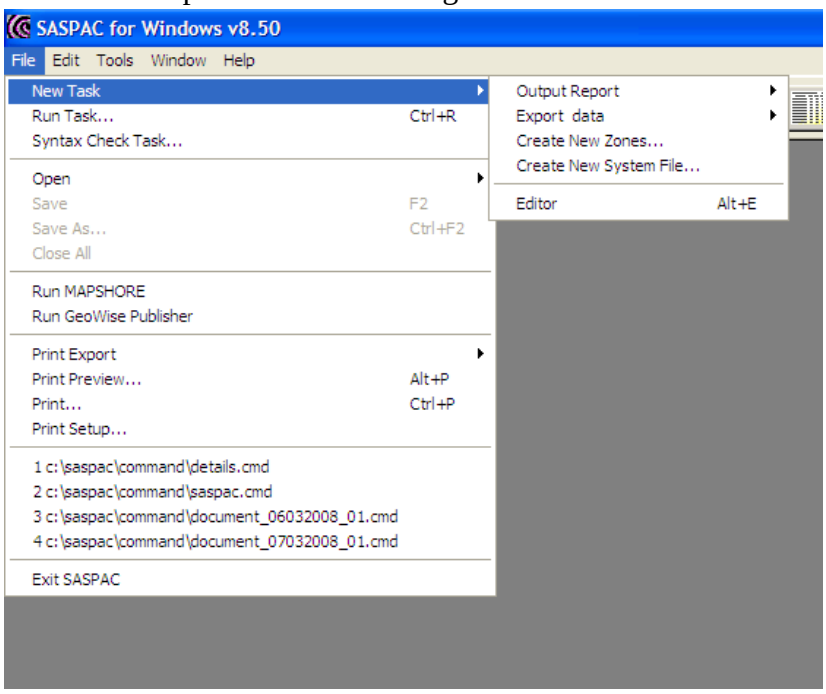


Figure 14: Selection of type of task to be undertaken.

Once the type of task has been specified, the answers to the other questions are then conveyed to SASPAC through the various windows and options that are presented to the user. Immediately following the selection of the type of task, the user is presented with the main task window. This window varies with the type of task selected, although there is a certain amount of similarity between them all.

Each main task window contains an area in which the input System Files are defined, and this area (on the left of the window) is the same for all tasks. The other areas of the main task window consist of a central area where, generally, manipulations and selections of the data are undertaken, and an area on the right which consists of buttons linked to other options related to the command file. Figure 15 shows the main task window for a ' Print Variables' task.

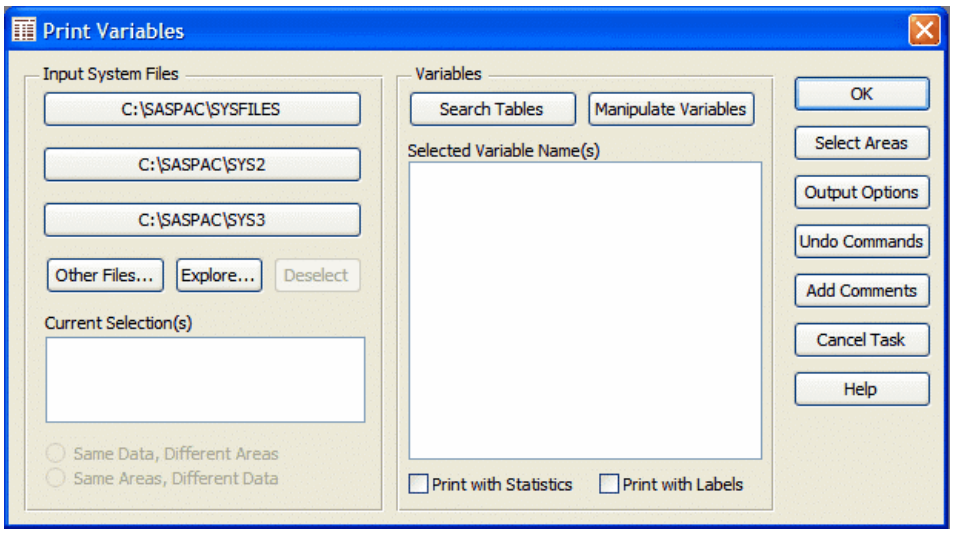


Figure 15: Main Task Window - for Print Variables..

There is (generally) no set order in which the elements of this window are completed, since the software will automatically order the generated statements into the required logical order. However, there will be instances when an input system file has to be defined before another element can be selected.

However, it should be noted that if the task is to access either the SWSC or SMS datasets, then these must be selected at the outset, as in these cases the windows used by SASPAC differ slightly from those presented to the user as a default, or when one of the other datasets is to be accessed.

What can be output from SASPAC?

The types of output available from SASPAC are:

"Printed" copy
- Variables
- Tables - can also be output as HTML files.
- Profiles

System Files

Interface Files
- Formatted
- Delimited
- CSV
- DIF
- WKS
- DBF

Each of these would be used as appropriate, and in this training document most will be covered. The examples shown in this manual refer to areas in Greater London, while the exercises to be undertaken during the training sessions will refer to areas of interest to the students.

Annex 1 Typical SASPAC.INI initialisation file for release 8.5.

```
[DDL FILES]
SAS91=C:\SASPAC\METADATA\SAS91.DDL
LBS91=C:\SASPAC\METADATA\LBS91.DDL
SAS81=C:\SASPAC\METADATA\SAS81.DDL
SWS81=C:\SASPAC\METADATA\SWS81.DDL

[HELP FILE]
HELP=C:\SASPAC\SASHELP.CHM

[GIS INFORMATION]
GISNAME= MAPSHORE
GISPROGRAMPATHNAME= C:\PROGRAM FILES\MAPSHORE\MAPSHORE.EXE
GISEXPORTTOSASPAC= C:\SASPAC\MAPPING
GISIMPORTFROMSASPAC= C:\SASPAC\INTERFAC
GISFINDWINDOW= INFOMAPWCLASS
GISPARAMETERS= /ONECOPY

[GEOWISE]
XMLSVGWARNING =
XMLPROGRAMPATHNAME = C:\SASPAC\GEOWISE\GWPUBLISHER.EXE
XMLPARAMETERS =
XMLFINDWINDOW = WINDOWSFORMS10.WINDOW.8.APP3
XMLIMPORTFROMSASPAC = C:\SASPAC\INTERFAC
XMLNAME = SASPAC GEOWISE WEB PUBLISHER


[FRAMEWORK, TABLE & CELL METADATA]
2001KS=ORIGIN:"SASPAC",FWK:"KS01.FWK",FDESC:"KEY
STATISTICS",YEAR:"2001",TYPE:"S",LOOKUP:"KS01DA",XML:"KS01"
2001UV=ORIGIN:"SASPAC",FWK:"UV01.FWK",FDESC:"UNIVARIATE
TABLES",YEAR:"2001",TYPE:"S",LOOKUP:"UV01DA",XML:"UV01"
2001ST=ORIGIN:"SASPAC",FWK:"ST01.FWK",FDESC:"STANDARD
TABLES",YEAR:"2001",TYPE:"S",LOOKUP:"ST01DA",XML:"ST01"
2001TT=ORIGIN:"SASPAC",FWK:"TT01.FWK",FDESC:"STANDARD THEME
TABLES",YEAR:"2001",TYPE:"S",LOOKUP:"TT01DA"
2001CS=ORIGIN:"SASPAC",FWK:"CS01.FWK",FDESC:"CENSUS AREA STATISTICS
TABLES",YEAR:"2001",TYPE:"S",LOOKUP:"CS01DA",XML:"CS01"
2001CT=ORIGIN:"SASPAC",FWK:"CT01.FWK",FDESC:"CENSUS AREA THEME
TABLES",YEAR:"2001",TYPE:"S",LOOKUP:"CT01DA"
2001KK=ORIGIN:"SASPAC",FWK:"KK01.FWK",FDESC:"VAR
TABLES",YEAR:"2001",TYPE:"K",LOOKUP:"KS01DA"
2001SWS=ORIGIN:"SASPAC",FWK:"SWS01.FWK",FDESC:"WORKPLACE-TRAVEL
TABLES",YEAR:"2001",TYPE:"D",LOOKUP:"SWS01DA"
2001SMS=ORIGIN:"SASPAC",FWK:"SMS01.FWK",FDESC:"MIGRATION
TABLES",YEAR:"2001",TYPE:"D",LOOKUP:"SMS01DA"
2001TVS=ORIGIN:"SASPAC",FWK:"TVS01.FWK",FDESC:"TRAVEL TO WORK/STUDY TABLES
(SCOTLAND)",YEAR:"2001",TYPE:"D",LOOKUP:"TVS01DA"
2001SCT=ORIGIN:"SASPAC",FWK:"SCT01.FWK",FDESC:"SPECIALLY-COMMISSIONED
TABLES",YEAR:"2001",TYPE:"H",LOOKUP:"SCT01DA"
2001AF=ORIGIN:"SASPAC",FWK:"AF01.FWK",FDESC:"ARMED FORCES
TABLES",YEAR:"2001",TYPE:"S",LOOKUP:"AF01DA"
1991SAS160=ORIGIN:"SASPAC",FWK:"TSAS160.FWK",FDESC:"SMALL AREA STATISTICS
(SAS)",YEAR:"1991",TYPE:"S",LOOKUP:"SASDA"
1991LBS160=ORIGIN:"SASPAC",FWK:"TLBS160.FWK",FDESC:"LOCAL BASE STATISTICS
(LBS)",YEAR:"1991",TYPE:"S",LOOKUP:"LBSDA"
1991SWS(A&B)=ORIGIN:"SASPAC",FWK:"TSWS.FWK",FDESC:"SPECIAL WORKPLACE
STATISTICS (A&B)",YEAR:"1991",TYPE:"S",LOOKUP:"SWS91DA"
```

1991SWS(C)=ORIGIN:"SASPAC",FWK:"SWSC91.FWK",FDESC:"SPECIAL WORKPLACE STATISTICS (C)",YEAR:"1991",TYPE:"D",LOOKUP:"SWSC91DA"
1991SMS=ORIGIN:"SASPAC",FWK:"SMS91.FWK",FDESC:"SPECIAL MIGRATION STATISTICS",YEAR:"1991",TYPE:"D",LOOKUP:"SMS91DA"
1991LRC=ORIGIN:"SASPAC",FWK:"LRC91.FWK",FDESC:"SPECIALLY-COMMISSIONED TABLES (LRC)",YEAR:"1991",TYPE:"S",LOOKUP:"LRCDA"
1991NI160=ORIGIN:"SASPAC",FWK:"NITAB160.FWK",FDESC:"NORTHERN IRELAND TABLES",YEAR:"1991",TYPE:"S",LOOKUP:"NIDA"
1981SASTABLE=ORIGIN:"SASPAC",FWK:"TABLE81.FWK",FDESC:"SMALL AREA STATISTICS (TABLES)",YEAR:"1981",TYPE:"S",LOOKUP:"81DA"
1981SASPAGE=ORIGIN:"SASPAC",FWK:"PAGE81.FWK",FDESC:"SMALL AREA STATISTICS (PAGES)",YEAR:"1981",TYPE:"S",LOOKUP:"81DA"
1981SWS=ORIGIN:"SASPAC",FWK:"SWS81TAB.FWK",FDESC:"SPECIAL WORKPLACE STATISTICS (A&B)",YEAR:"1981",TYPE:"S"
1981JUVOS=ORIGIN:"SASPAC",FWK:"JUVOS.FWK",FDESC:"JUVOS",YEAR:"1981",TYPE:"S"

[CONFIGURATION FILE]
CFG = 8.50

[USER CONFIGURATION]
CFG= C:\DOCUME~1\ALEWIS\LOCALS~1\TEMP\SASPAC.CFG
SYSTEM1 FILE DIRECTORY = J:\TEAM FOLDERS\SASPAC\SASPAC_DATA\2001 CENSUS RATIONALISED AND CONSOLIDATED SYSTEM FILES (CURRENT)
SYSTEM2 FILE DIRECTORY = C:\SASPAC\SYS2
SYSTEM3 FILE DIRECTORY = C:\SASPAC\SYS3
COMMAND FILE DIRECTORY NAME = C:\SASPAC\COMMAND
REPORT FILE DIRECTORY NAME = C:\SASPAC\REPORT
LOG FILE DIRECTORY NAME = C:\SASPAC\LOG
INTERFACE FILE DIRECTORY NAME = C:\SASPAC\INTERFAC
FRAME DATA FILE DIRECTORY NAME = C:\SASPAC\FRWDATA
OPCS DATA FILE DIRECTORY NAME = C:\SASPAC\OPCS
FRAMEWORK FILE DIRECTORY NAME = C:\SASPAC\FRWORK
SASPAC4 FILE DIRECTORY NAME = C:\SASPAC\SASPAC4
CONVERT FRAMEWORK BORDER = YES
KEY FOR VAR IN FRAMEWORK = @
KEY FOR TITLE IN FRAMEWORK = &
PAGE LENGTH = 80
PAGE WIDTH = 80
DATA FORMAT OF FRAMEWORK = ******
POSITION DEFAULT OF FRAMEWORK = R
MAXIMUM PAGE = 200
PAGE SYMBOL FOR REPORT = $
TEXT QUOTE FOR DELIMITED FILE = '
DELIMITER FOR DELIMITED FILE = ;
MISSING VALUE SYMBOL = MISS
RECORDS FILE DIRECTORY NAME = C:\SASPAC\RECORDS
HTML FILE DIRECTORY = C:\SASPAC\HTML
MAPPING GAZETTEER FILE DIRECTORY = C:\SASPAC\MAPPING

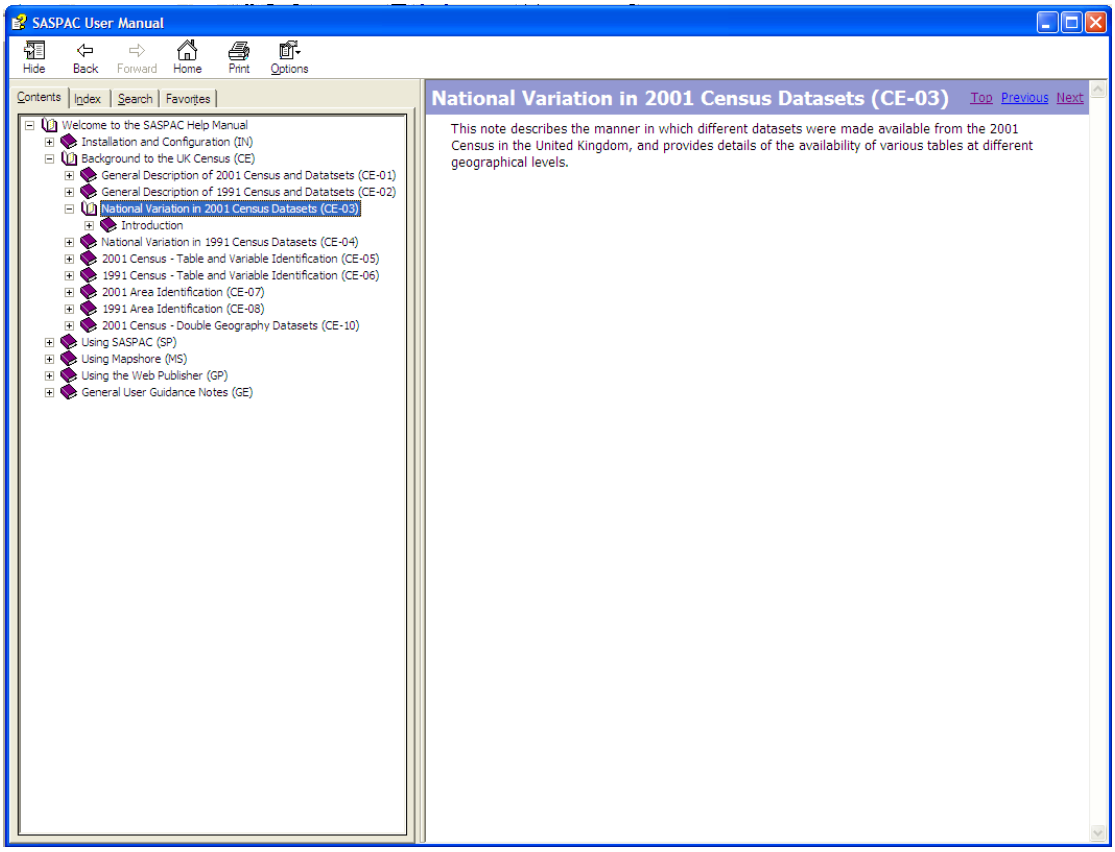
[FILES]
4
C:\SASPAC\COMMAND\DETAILS.CMD
C:\SASPAC\COMMAND\SASPAC.CMD
C:\SASPAC\COMMAND\DOCUMENT_06032008_01.CMD
C:\SASPAC\COMMAND\DOCUMENT_07032008_01.CMD

[SETTINGS]
ENHANCE INCLUDE/EXCLUDE READABILITY = 0
WARN IF MISSING XFF OR XVM = 0
LARGE TOOLBAR BUTTONS = 1
800X600 WARNING = 0

REZONECALCFILE = C:\SASPAC\METADATA\KS PER_NUM ID.TXT
COMMANDLINEMAX = 140
DEFAULT METADATA FOLDER = C:\SASPAC\METADATA\
XMLHEADERFILE = C:\SASPAC\METADATA\XMLHEAD.TXT
EDITING FONT = COURIER NEW,-12
PRINTING FONT = COURIER NEW
INTERNET BROWSER = "INTERNET EXPLORER","C:\PROGRAM FILES\INTERNET EXPLORER\IEXPLORE.EXE"
GMSIZE = 550,550
GMAPIKEY =

[PRINTING]
DIAGS=OFF
PAPER = A4
WIDTH = 210.00
HEIGHT = 297.00
LM = 12.00
RM = 12.00
TM = 12.00
BM = 12.00

Annex 2 Valid National Table Identifiers are detailed in SASPAC's User Manual, which can be accessed through the Help menu in the software. Census User Guidance note 'CE03' contains a series of tables that identify which datasets are available for different geographies and countries.


SASPAC User Manual
Hide
Back
Forward
Home
Print
Options
Contents
Index
Search
Favorites
Welcome to the SASPAC Help Manual
Installation and Configuration (IN)
Background to the UK Census (CE)
General Description of 2001 Census and Datatsets (CE-01)
General Description of 1991 Census and Datatsets (CE-02)
National Variation in 2001 Census Datasets (CE-03)
Introduction
National Variation in 1991 Census Datasets (CE-04)
2001 Census - Table and Variable Identification (CE-05)
1991 Census - Table and Variable Identification (CE-06)
2001 Area Identification (CE-07)
1991 Area Identification (CE-08)
2001 Census - Double Geography Datasets (CE-10)
Using SASPAC (SP)
Using Mapshore (MS)
Using the Web Publisher (GP)
General User Guidance Notes (GE)
National Variation in 2001 Census Datasets (CE-03)
Top
Previous
Next
This note describes the manner in which different datasets were made available from the 2001 Census in the United Kingdom, and provides details of the availability of various tables at different geographical levels.

Module 1.

Selection of Variables.

In the main task window - whatever the task - there will be a Windows Command button that allows access to various searching mechanisms for identification of variables required by users. This also enables creation of the variable identities through on-screen identification. The 'Print Variables' main window as shown in Figure 15 on page 18 has this button in the central area with the text 'Search Tables'. Use of this button leads to a window with a drop down menu and four buttons leading to different Property Sheets as shown in Figure 16.

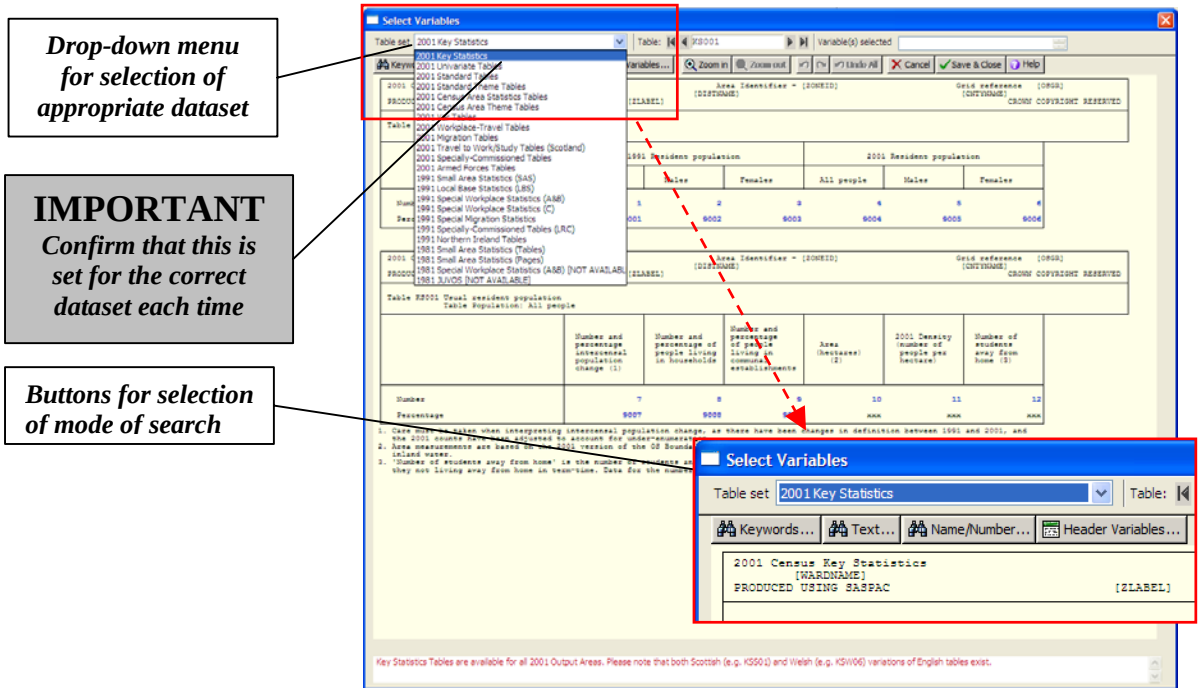


Figure 16: Selection of dataset to be searched.

Since the datasets that can be accessed through SASPAC contain different data and have different cell identifiers, the drop down menu is used to tell SASPAC which dataset is being accessed This is done by the user clicking on the required one.

The four buttons on the 'Select Variables' window are:

Keywords: SASPAC contains a pre-defined dictionary of the more common keywords or topics found in the SAS and LBS datasets. This dictionary may be searched when this Tab (which is the default) is selected. One or more items may be selected in the left-hand 'Keywords' window, and transferred to the right-hand 'Keywords' window through use of the '**Add**>>' button. Once the keywords are transferred, the bottom right-hand text box will contain a list of the tables which, according to the dictionary, refer to that keyword. Unnecessary keywords may be removed through use of the '**<<Remove**' button, following highlighting. Clicking on one of these table numbers will activate the '**Table n**' Tab and display the table on screen.

Text: If a keyword is not found in the dictionary, or the dictionary references are incomplete, this Tab may be used to initiate a full text search of the table layouts to find occurrences of any text string. For example, the word "farmers" may be entered in the appropriate box to discover that tables 82, 86 and 89 of the SAS contain the word "farmers".

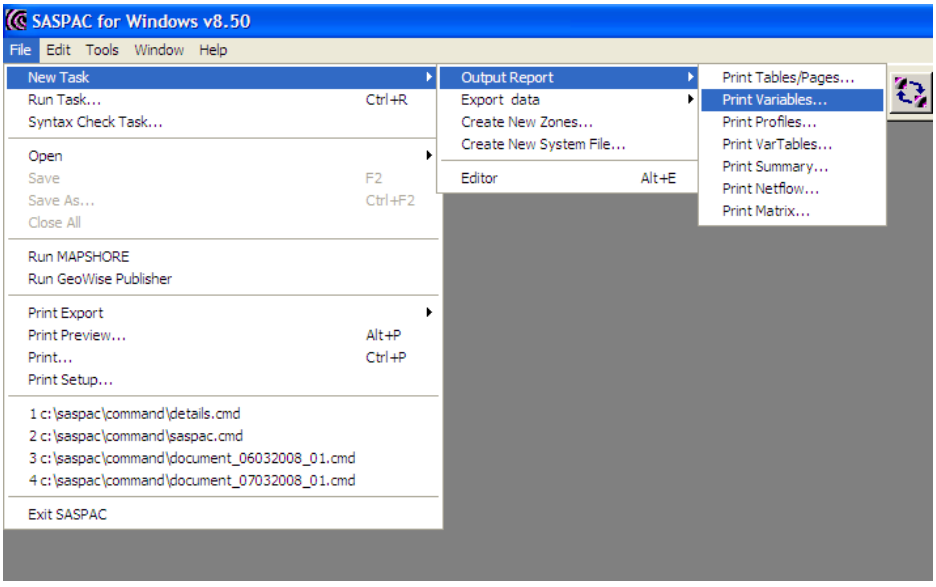
Name/Number: If the approximate title of the table is known, this tab may be used to show a list of all tables in the selected dataset, from which the required table is selected. The selected table is then shown on screen.

Header Variables: In addition to the count variables that are available through the tables, there is a set of 'Header variables' as defined in pages A2.2 to A2.4 of the SASPAC for Windows User Manual. If any of these are required for use in a task, this Tab can be used to access them.

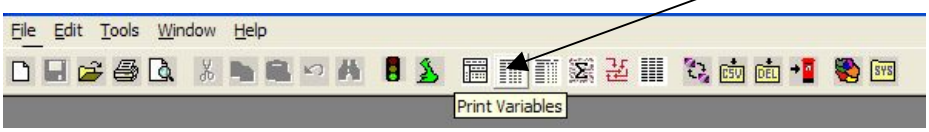
Once the required table is identified, the layout on screen may then be used to identify the required variable(s). When the variables are selected, their full identities are written into the text box at the top of the window. Variable selection may be undertaken, either by a point and click method or by the 'rubber-band' method. In the latter, the top left hand corner of a bounding rectangle is chosen by clicking on it, and then by dragging to the bottom right hand corner. The following example shows the use of this.

Example: From the 2001 Census, for all wards in the London Borough of Enfield print the total of usual residents born in Scotland, the total number of widowed females aged 70 to 74, and aged 75 to 79, the total number of households with 4 or more cars or vans available, and the number of males employed in the mining industry. Include the ward name in the output.

1. The first question to ask is - "What sort of task is this?" and the answer is that it is to produce variables as a printed report. The first stage is therefore to select the appropriate menu items for defining such a task. This is done as shown in the following image.

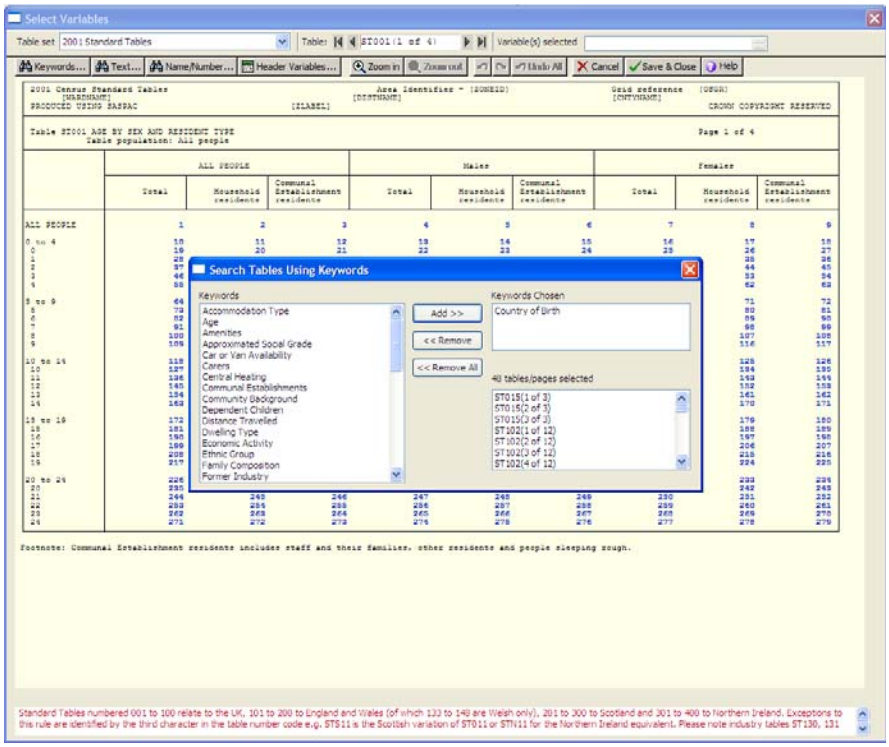


Alternatively, instead of using the menu route, the user now (as of release 6.5) can make use of the button bar to select certain tasks. The ‘Print Variables’ button is shown in this image.

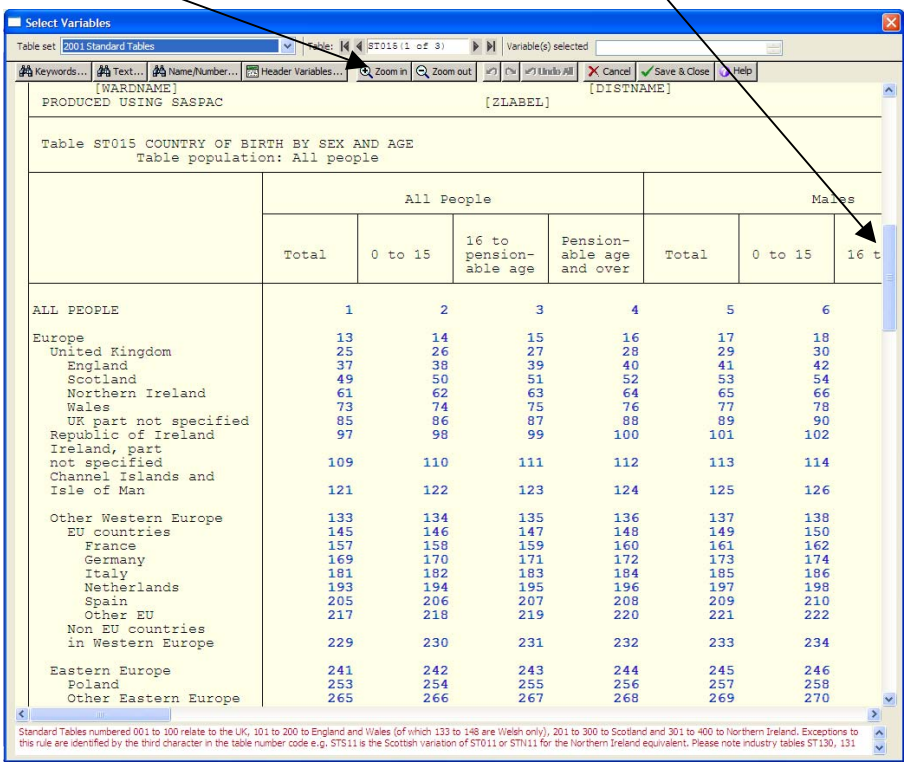


2. Following either of these selections, the main 'Print Variables' task window is opened. At this point, we can choose the input System Files, and use the 'Search Tables' button to select the required variables. There is no mandatory order in which these operations are done. Since we may not be certain of whether or not all the data are in the same System File, we will initially search for the variables by using the 'Search Tables' button. For the first variable - residents born in Scotland - we will use the '**Keywords…**' button and as shown in the next screen capture, use the keyword 'Country of Birth'.

   We must first of all remember that as we are looking for ward level data, it is advisable, but not mandatory to look in the Standard Tables and must therefore select that dataset from the drop-down list. The dictionary now indicates that tables ST015, ST102, ST133, ST150, ST202, ST236, ST314 and ST373 all include this keyword. However, bearing in mind what was stated on page 11, we can immediately disregard the last four as they only apply to areas in Scotland and Northern Ireland. Note that tables that span more than one page have multiple entries in this list.



3. Clicking on the first table number in the right hand box and closing the '**Search Tables Using Keywords**' box, will open the table as a layout. Use of the '**Zoom in**' button three times together with scrolling down will bring up the window shown here.

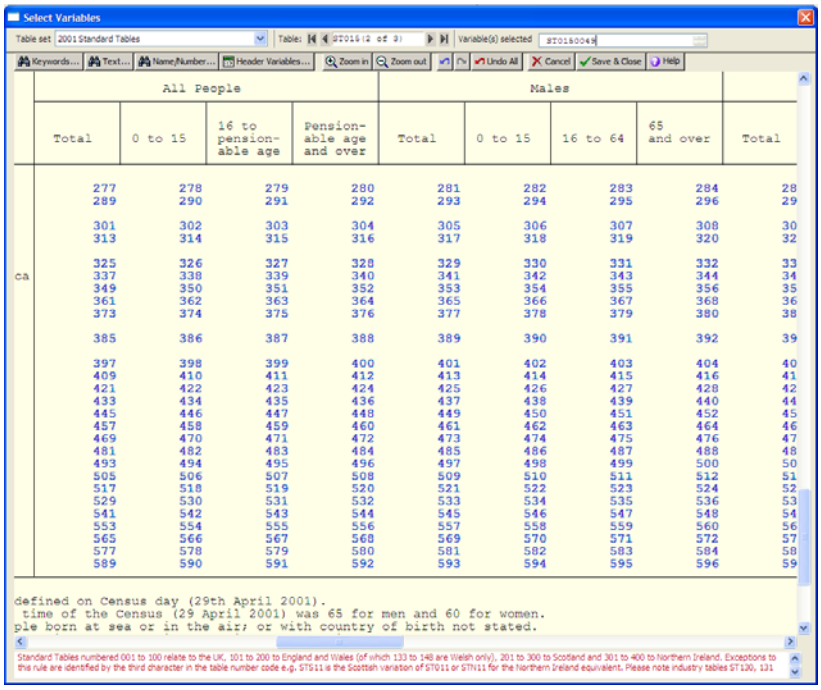

Select Variables
Table set 2001 Standard Tables
Table: ST015 (1 of 3)
Variable(s) selected
Keywords...
Text...
Name/Number...
Header Variables...
Zoom in
Zoom out
Undo All
Cancel
Save & Close
Help
[WARDNAME]
[DISTNAME]
PRODUCED USING SASPAC
[ZLABEL]
Table ST015 COUNTRY OF BIRTH BY SEX AND AGE
Table population: All people
All People
Males
Total
0 to 15
16 to pension-able age
Pension-able age and over
Total
0 to 15
16 t
ALL PEOPLE 1 2 3 4 5 6
Europe 13 14 15 16 17 18
United Kingdom 25 26 27 28 29 30
England 37 38 39 40 41 42
Scotland 49 50 51 52 53 54
Northern Ireland 61 62 63 64 65 66
Wales 73 74 75 76 77 78
UK part not specified 85 86 87 88 89 90
Republic of Ireland 97 98 99 100 101 102
Ireland, part not specified 109 110 111 112 113 114
Channel Islands and Isle of Man 121 122 123 124 125 126
Other Western Europe 133 134 135 136 137 138
EU countries 145 146 147 148 149 150
France 157 158 159 160 161 162
Germany 169 170 171 172 173 174
Italy 181 182 183 184 185 186
Netherlands 193 194 195 196 197 198
Spain 205 206 207 208 209 210
Other EU 217 218 219 220 221 222
Non EU countries in Western Europe 229 230 231 232 233 234
Eastern Europe 241 242 243 244 245 246
Poland 253 254 255 256 257 258
Other Eastern Europe 265 266 267 268 269 270
Standard Tables numbered 001 to 100 relate to the UK, 101 to 200 to England and Wales (of which 133 to 148 are Welsh only), 201 to 300 to Scotland and 301 to 400 to Northern Ireland. Exceptions to this rule are identified by the third character in the table number code e.g. STS11 is the Scottish variation of ST011 or STN11 for the Northern Ireland equivalent. Please note industry tables ST130, 131

This shows that the variable identifying the number of persons born in Scotland is the forty ninth in the table and clicking on this produces a red tick, and the variable identity is written to the text box at the top of the window. Note that clicking again on the cell, will remove both the red tick and the entry in the text box. In addition, the entry in the text box may be deleted manually, which also removes the tick.

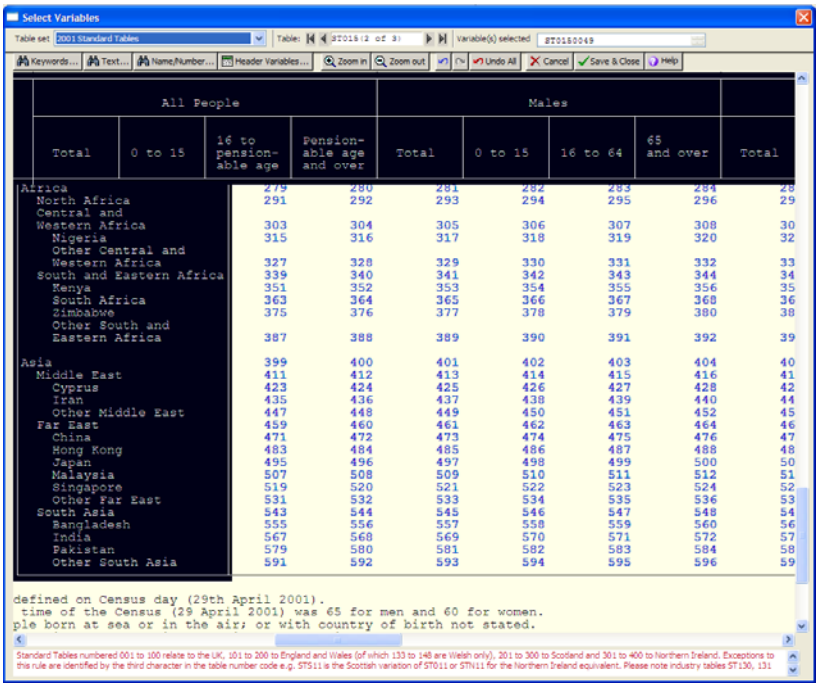
4. The next two variables required are found by a similar method using 'Age', 'Marital Status', and 'Sex' as the keywords (remembering to remove 'Country of Birth' from the right hand **'Keywords'** box) which will bring up Table ST002 as the only table containing these three keywords.

   Firstly, we note that this table is split over two tables, and the first page refers only to Males. To reach the page showing the Females, we can use the next page arrow button (located next to the table number above the layout). When that is done, to find the desired variables, the 'Zoom in' button may again be used. Using the scroll bars to move to the part of the table containing the appropriate counts will produce a window that looks similar to that shown below.

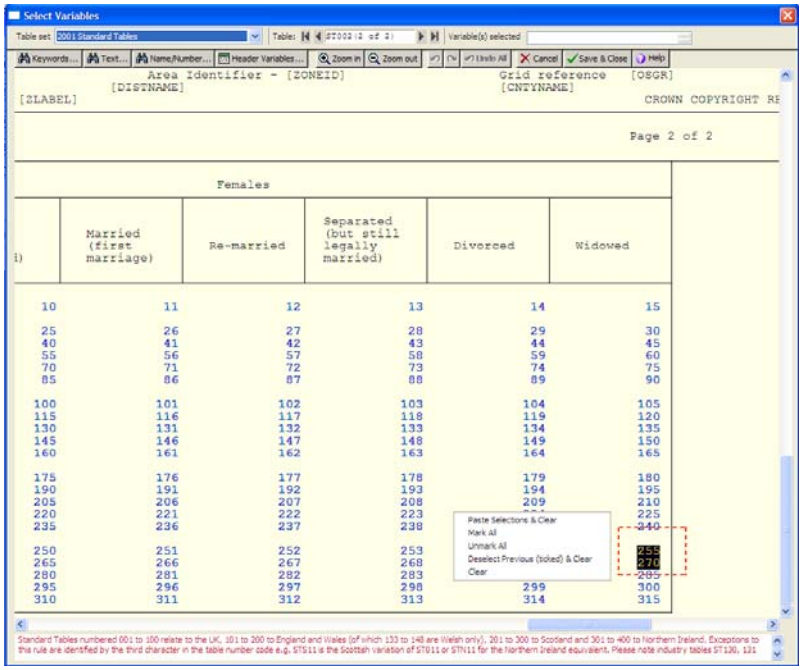
   By scrolling and zooming, we have lost all the useful information surrounding the table, and selection of the correct counts or variables is made very difficult.


Select Variables
All People
Males
Total
0 to 15
16 to pension-able age
Pension-able age and over
Total
0 to 15
16 to 64
65 and over
Total
defined on Census day (29th April 2001).
time of the Census (29 April 2001) was 65 for men and 60 for women.
ple born at sea or in the air; or with country of birth not stated.

5. To overcome this problem, we right click (i.e. click the right mouse button) anywhere in the body of the table shown in above. This has the effect of replicating (on a black background), the text that appears directly above and to the left of the cursor. This is shown in the next screen image.



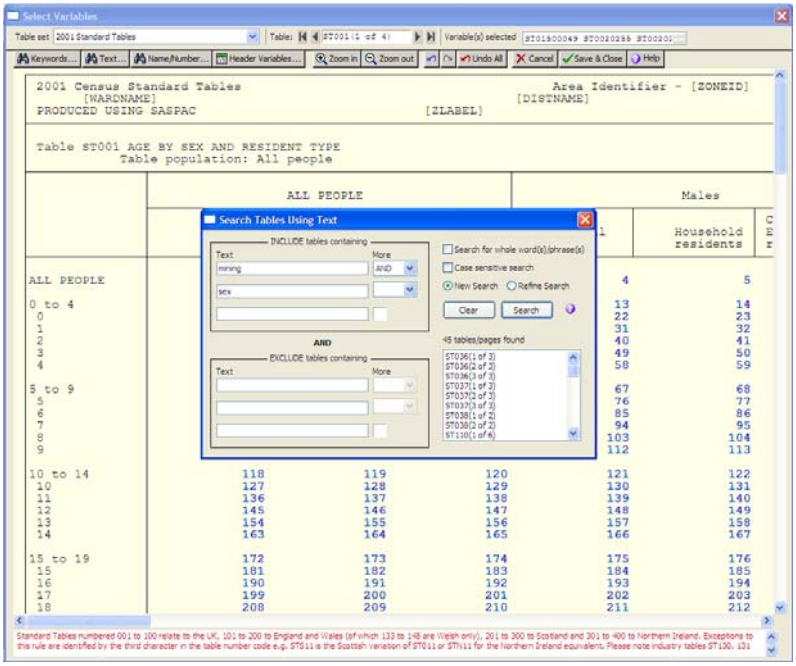
6. The presence of the row and column headings now allows us to select the appropriate counts. As the two that are required are adjacent, we can use the 'rubber-banding ' technique that makes use of the right mouse button again. First the left button is clicked and a drag operation is used to draw a rectangle around the required cells. The right mouse button is then used to bring up a menu as shown below.



7. Use of the 'Paste Selection and clear will transfer the contents of the 'rubber-band' to the text box at the top of the window. The selected counts will be shown in the window with a red tick. As with individually selected variables, these may be deselected by clicking on them, using the ‘Undo’ button or by editing the text box.

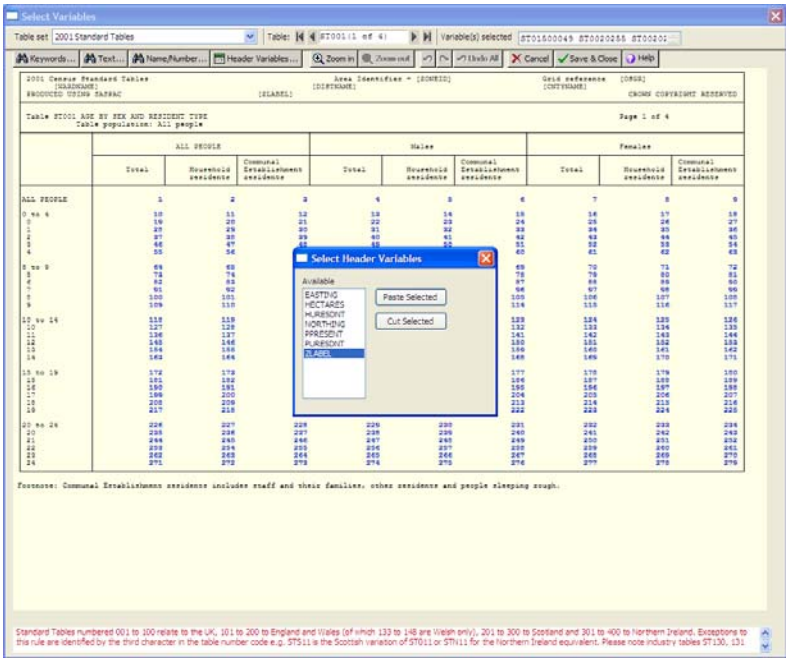
8. The procedure is then repeated to select the next variable, using "Car availability as the keyword to identify its code. Remember again that before each search is undertaken, the keyword currently in the right hand box has to be removed, and the other variables then selected.

   An alternative to searching by keyword which is especially useful when the appropriate keyword is not known, is to use the ‘**Text…**’ button in the ‘**Select Variables**’ window as is shown in the following image. It can be seen that this window allows for multiple searches on the text contained within the tables, but it is usually best to keep the searching criteria as simple as possible.

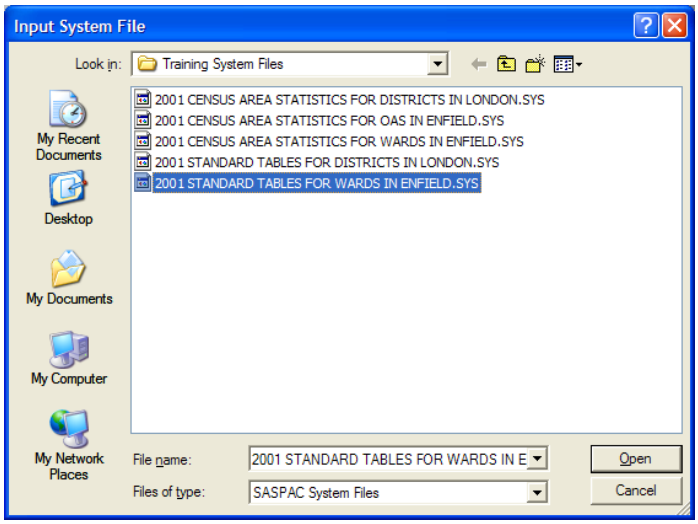


9. After searching the tables for the text ‘mining’ and ‘sex’ a set of tables containing this text is offered to the user. The appropriate variable is now selected from table ST036. The five variables which we have identified are ST0150049 ST0020255 ST0020270 ST0590031 ST0360229, and having selected these, we could return to select the appropriate input System Files. To return to the 'Print Variables' main task window, we use the 'Save & Close' button below the Variable(s) selected window.

   However, we are also asked to output the area name, and this is obtained through use of the ‘**Header Variables…**’ button in the ‘Print Variables’ window, as shown below. This window may be accessed again even though we have returned to the main task window, and this shows that there is no pre-defined order in which the windows/facilities may be accessed. On the left of this window can be seen all the header variables that are available with printing variables, and the one that we require is ‘ZLABEL’. This is selected in the left hand box, and then the ‘Paste Selected’ button is used to transfer it to the variable(s) selected box. Note that whenever the header variable is selected, it will appear at the beginning of the list.



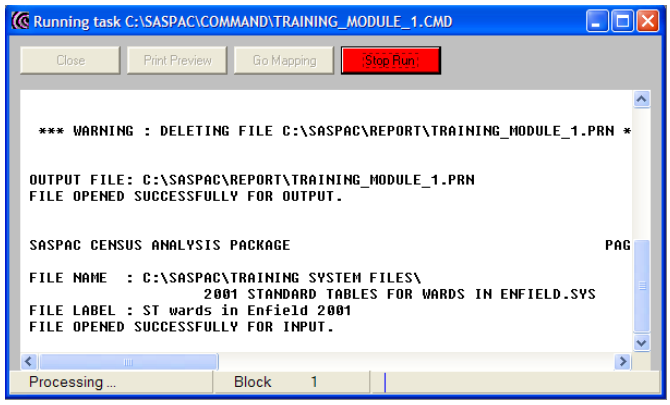
10. The '**Print Variables**' main task window presents the user with up to three Command Buttons linked to directories as defined in the user's configuration. If the required system file is not in one of these directories, then the standard Windows browse facilities are available through the '**Other Files…**' Command Button. In this example, the files we want are in the second default directory, and using the appropriate button opens a standard Windows dialogue box as shown in the next screen image.



11. The command file is now complete and we are ready to run the task. This is achieved by using the '**OK**' Command Button, which generates the '**Task**' window. In this window, the user can either run the command file under the default name of SASPAC.CMD, or allocate a new name under which the command file will be saved. If a user-defined name is given to the command file, the output report or print file will default to this name as well, unless the report file is forced to a user-defined name as well.

    Providing a user-defined name is good practice as this will obviate the possibility of overwriting an existing command file, and also allows for it to be used later as a template which can be edited.

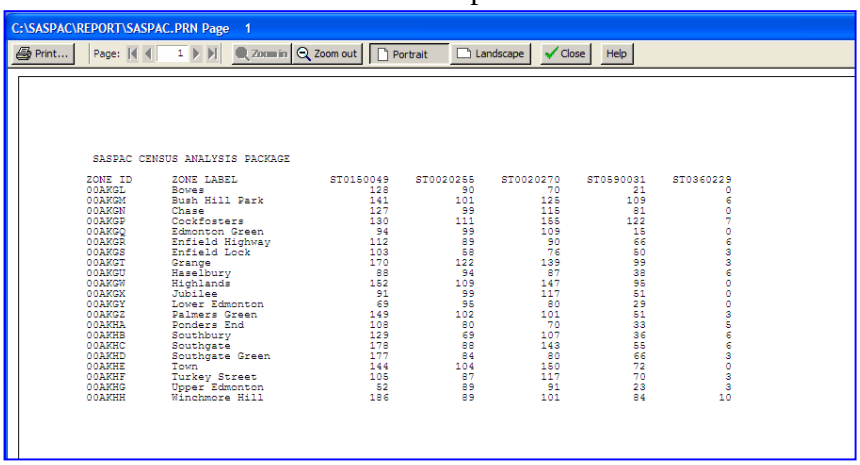
12. When we are satisfied that all is as required, the '**OK**' Command Button is used to run the task. Otherwise the '**Cancel**' Command Button may be used to return to the 'Print Variables' window to make further selections or amendments. When the task is run, the 'Running Task' window echoes the command file that has been generated and reports any warning or error messages generated by the task.

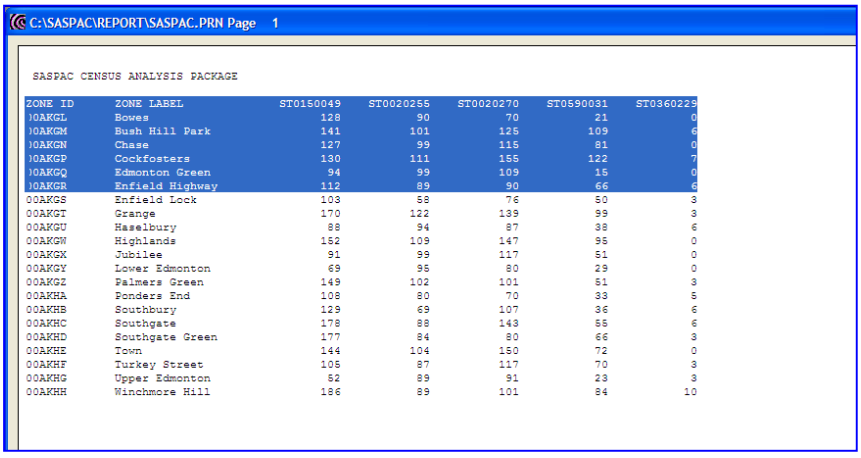

Running task C:\SASPAC\COMMAND\TRAINING_MODULE_1.CMD
Close
Print Preview
Go Mapping
Stop Run
*** WARNING : DELETING FILE C:\SASPAC\REPORT\TRAINING_MODULE_1.PRN *
OUTPUT FILE: C:\SASPAC\REPORT\TRAINING_MODULE_1.PRN
FILE OPENED SUCCESSFULLY FOR OUTPUT.
SASPAC CENSUS ANALYSIS PACKAGE PAG
FILE NAME : C:\SASPAC\TRAINING SYSTEM FILES\
2001 STANDARD TABLES FOR WARDS IN ENFIELD.SYS
FILE LABEL : ST wards in Enfield 2001
FILE OPENED SUCCESSFULLY FOR INPUT.
Processing ...
Block 1

14. The command file generated by the above sequence is:

```
input system file name = "C:\SASPAC\Training /
System Files\2001 STANDARD TABLES FOR WARDS IN /
ENFIELD.SYS"
print variables ZLABEL ST0150049 ST0020255 /
ST0020270 ST0590031 ST0360229
output print file = /
C:\SASPAC\REPORT\TRAINING_MODULE_1.PRN
end
finish
```

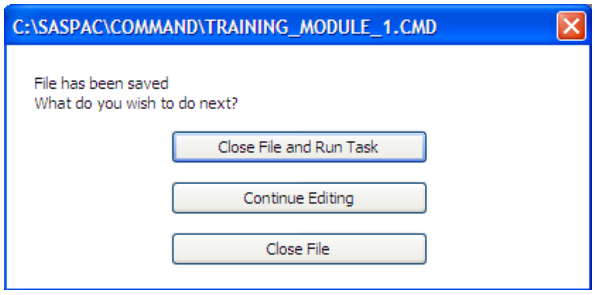
15. When the task is complete, the **'Print Preview'** Command Button may be used to examine the output. The effect of this is to display the output as shown in the first image below. The '**Print Preview'** option is also available through the '**File**' menu option. Also available through this option is the '**Print Export/Output Text File…**' facility. This would display the output as shown in the second image below. The difference between the two is that the first is a graphic (or picture) image that cannot be edited in any way, while the second is simply text which can be edited or 'cut and pasted' into another document.


C:\SASPAC\REPORT\SASPAC.PRN Page 1
Print...
Page: 1
Zoom in
Zoom out
Portrait
Landscape
Close
Help
SASPAC CENSUS ANALYSIS PACKAGE
ZONE_ID ZONE LABEL ST0150049 ST0020255 ST0020270 ST0590031 ST0360229
00AKGL Bowes 128 90 70 21 0
00AKGM Bush Hill Park 141 101 125 109 6
00AKGN Chase 127 99 115 81 0
00AKGP Cockfosters 130 111 155 122 7
00AKGQ Edmonton Green 94 99 109 15 0
00AKGR Enfield Highway 112 89 90 66 6
00AKGS Enfield Lock 103 88 76 50 3
00AKGT Grange 170 122 139 99 3
00AKGU Haselbury 88 94 87 38 6
00AKGW Highlands 152 109 147 95 0
00AKGX Jubilee 91 99 117 51 0
00AKGY Lower Edmonton 69 95 80 29 0
00AKGZ Palmers Green 149 102 101 51 3
00AKHA Ponders End 108 80 70 33 5
00AKHB Southbury 129 69 107 36 6
00AKHC Southgate 178 88 143 55 6
00AKHD Southgate Green 177 84 80 66 3
00AKHE Town 144 104 150 72 0
00AKHF Turkey Street 105 87 117 70 3
00AKHG Upper Edmonton 52 89 91 23 3
00AKHH Winchmore Hill 186 89 101 84 10

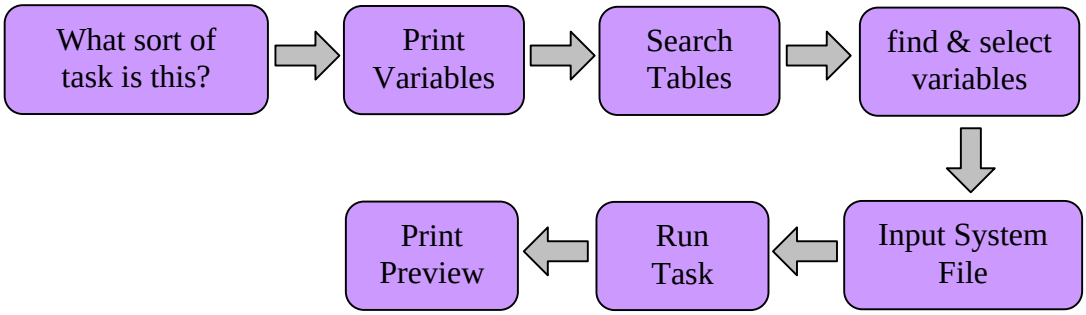

C:\SASPAC\REPORT\SASPAC.PRN Page 1
SASPAC CENSUS ANALYSIS PACKAGE
ZONE_ID ZONE LABEL ST0150049 ST0020255 ST0020270 ST0590031 ST0360229
00AKGL Bowes 128 90 70 21 0
00AKGM Bush Hill Park 141 101 125 109 6
00AKGN Chase 127 99 115 81 0
00AKGP Cockfosters 130 111 155 122 7
00AKGQ Edmonton Green 94 99 109 15 0
00AKGR Enfield Highway 112 89 90 66 6
00AKGS Enfield Lock 103 88 76 50 3
00AKGT Grange 170 122 139 99 3
00AKGU Haselbury 88 94 87 38 6
00AKGW Highlands 152 109 147 95 0
00AKGX Jubilee 91 99 117 51 0
00AKGY Lower Edmonton 69 95 80 29 0
00AKGZ Palmers Green 149 102 101 51 3
00AKHA Ponders End 108 80 70 33 5
00AKHB Southbury 129 69 107 36 6
00AKHC Southgate 178 88 143 55 6
00AKHD Southgate Green 177 84 80 66 3
00AKHE Town 144 104 150 72 0
00AKHF Turkey Street 105 87 117 70 3
00AKHG Upper Edmonton 52 89 91 23 3
00AKHH Winchmore Hill 186 89 101 84 10

16. The command file created in this task may, if required, be edited to produce different output. This depends upon the user having sufficient information to be able to provide the correct information without using the facilities of the menu system. The command file is opened for editing either through use of the 'most recent files' facility under the '**File'** menu item, or through the

'**File/Open/Open Command File…**'. After editing, the '**File/Save'** menu option is used , the result of which is to present the user with the following window in which the most appropriate selection is made. If the command file has been edited, but no change made to the name of the output Print file, the user will be asked if they wish to overwrite an existing file. If the answer to this 'No', then the command file will be saved, but not run.



17. Summary of task sequence:



*make sure the system file contains the geographical areas and variables you are interested in

Module 2.

Selection of Tables.

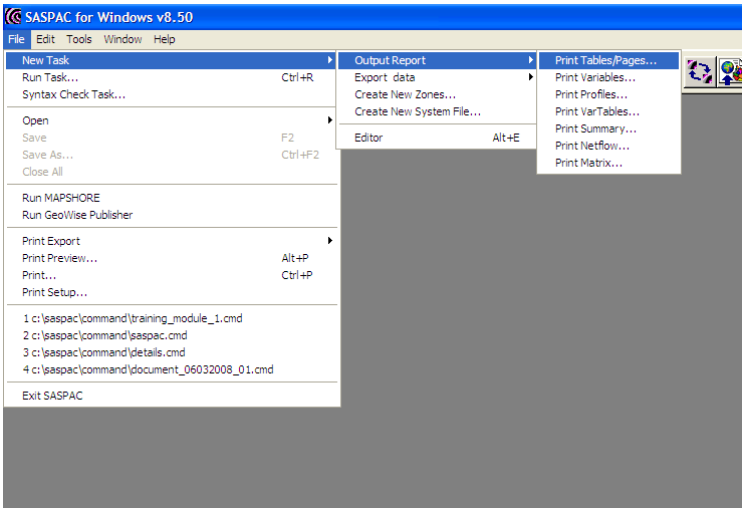
When selecting tables, the main task window has an area where the user is asked to identify the text file which is to be merged with the data file(s) to produce the appropriate tables. It is important that the framework file chosen matches the data type of the data file. For example, if the input system file contains Standard Table data, then the framework file chosen must be the Standard Tables File.

The standard procedure for selection of tables is similar to that for the selection of variables, with access to the searching facilities being made through a '**Select Tables**' Command Button rather than a '**Search Tables**' Command Button. When this button is used, the window offered to the user mirrors that presented while searching for variables, except there are only three buttons - the Header Variables button is not present.

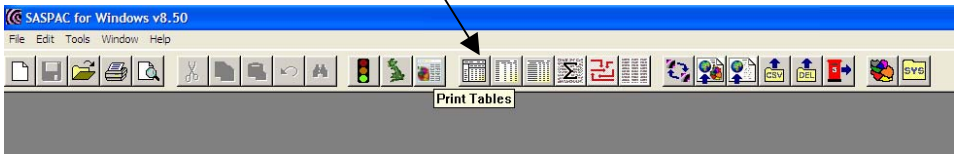
However, there is one major difference between the two procedures. In variable selection, clicking the mouse button on a variable identifier automatically writes that code to the 'Variable(s) selected' text box. In table selection, a table identified as required is only written to the text box when the '**Select Table**' Command Button is clicked in the '**Table(s) Selected**' window.

Example: For all London Boroughs print the table which shows the Tenure and Persons per Room, and also the one which shows numbers of Buddhists by Sex and Age. Convert the tables to HTML format and input a table into Excel.

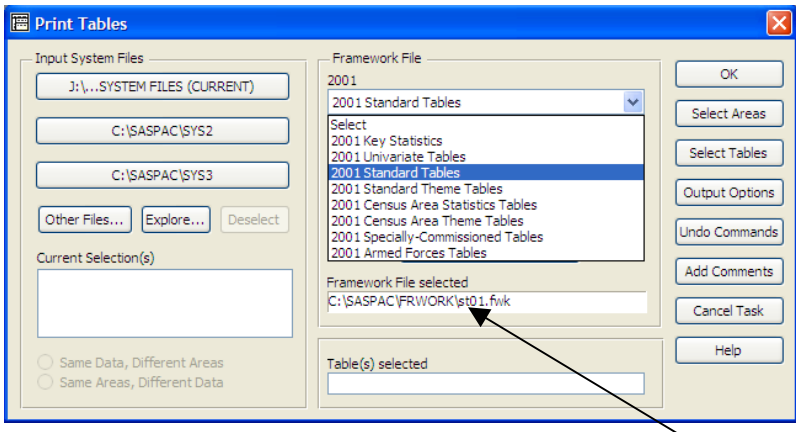
1. Again the question is "What sort of task is this?" to which the answer is that it is to produce tables as a printed report, and so the first stage is to identify this to SASPAC through the menu options as shown on the following screen.



As with the 'Print Variables' task discussed in Module 1, there is now a shortcut to the main 'Print Tables' task window, available through the Button Bar, and this is shown below.

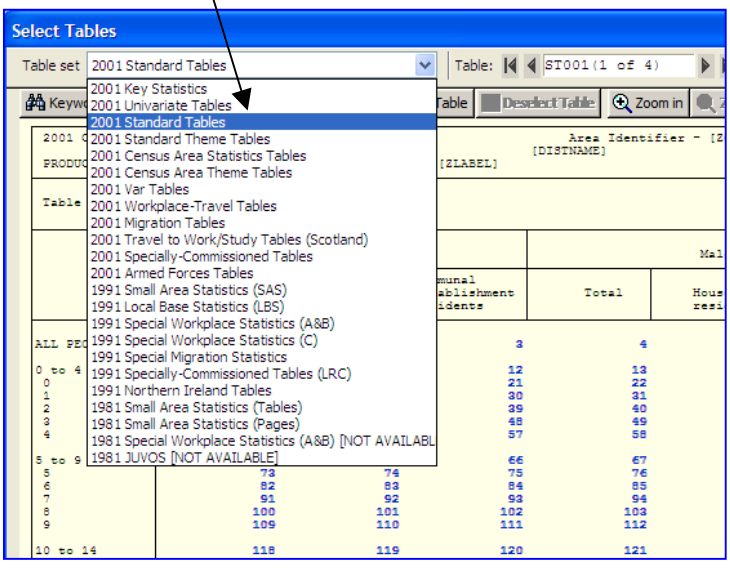


2. In the subsequent '**Print Tables**' window we can do three things - select the input system file(s), select the appropriate framework file, or select the required table(s). As with the '**Print Variables**' task, there is no set order in which these options must be done, and it is sometimes appropriate to leave selection of the system files until the tables have been identified. ***(This is not the case if double geography datasets are to be accessed.)*** As we are looking at London Borough (or district) level data, the dataset that we shall use is the Standard Tables, and so the appropriate framework file can be selected as the from the drop down 2001 list. If other data, 1991 for example, were required, selection would be made from the drop down 1991 list.



Note that when the framework file is selected, its full path name appears in the text box below the drop down list.

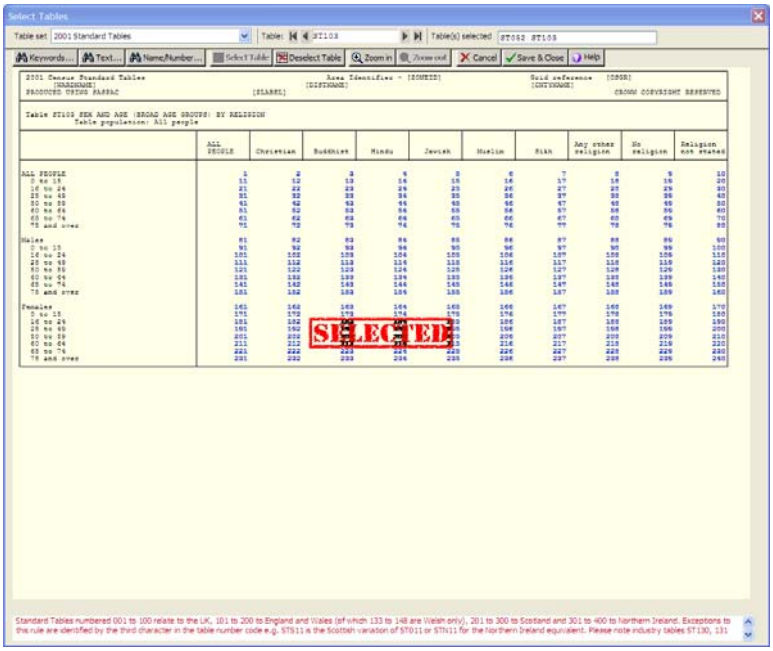
3. The '**Select Tables**' Command Button can now be used to enter the searching options. The first action is to select the appropriate dataset to search, which in this case is the 2001 Standard Tables.



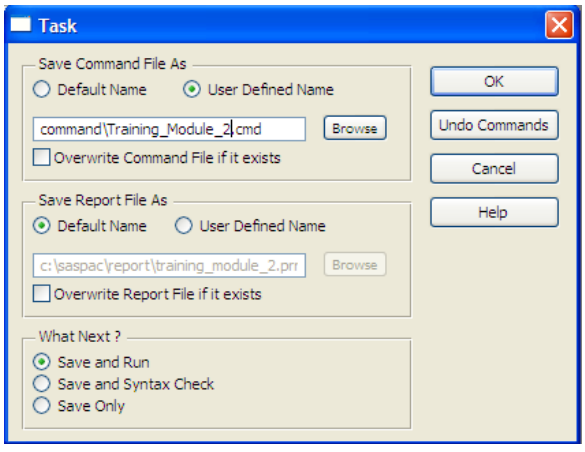
As in selecting variables, keywords or text can be searched, and by using the former we can ascertain that table ST052 is the required table for tenure and persons per room. To find the required table for Buddhists, we shall use the '**Text…**' button and enter the text '**Buddhist**' in the '**INCLUDE**' text box.

This shows us that there are several tables containing this word in the text, and if the first is selected (table ST103) table layout is offered to us. This shows that age and sex are also constituents of this table as detailed in the image shown on the next page.

When the '**Select Table**' Command Button is used, the table identifier is written to the '**Table(s) selected**' text box, and the table layout is stamped 'SELECTED'.



4. Once the tables required for printing have been identified, there still remains a requirement to identify the input system files. The file required this time is **2001 STANDARD TABLES FOR DISTRICTS IN LONDON.SYS**, as the task is to print tables from the Standard Tables for all London Boroughs. After this file is selected, the task may be run.



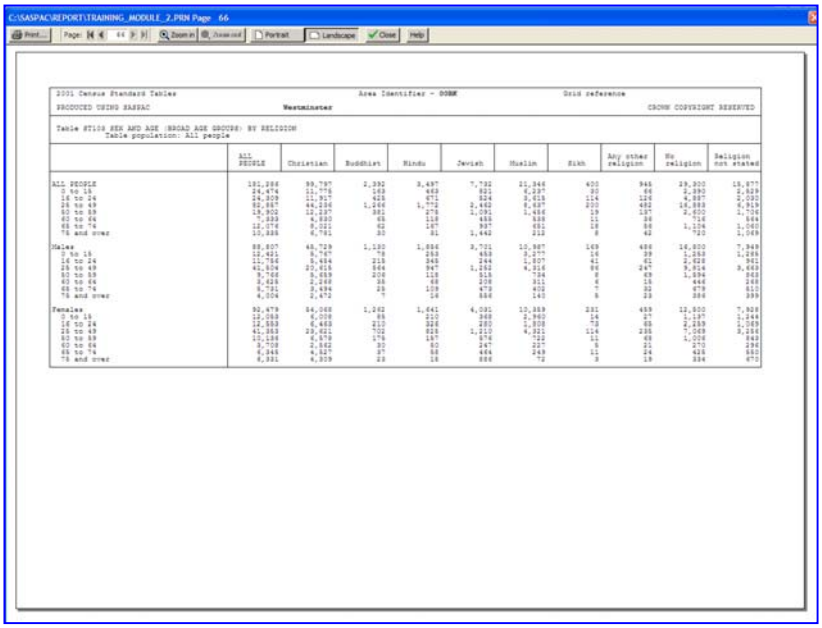
Every time a new task is created through the appropriate menu/window sequence, the last stage is the '**Task**' window as shown above. This window allows the user to allocate a Command File name other than the default of SASPAC.CMD. To do this, the '**User Defined Name**' radio button is selected, and an appropriate name is then entered in the text box. There is no

obligation to change from the default, but if there is a need to save any command files or output report files, then it is advisable that an appropriate name is allocated.

5. The command file generated by this sequence is:

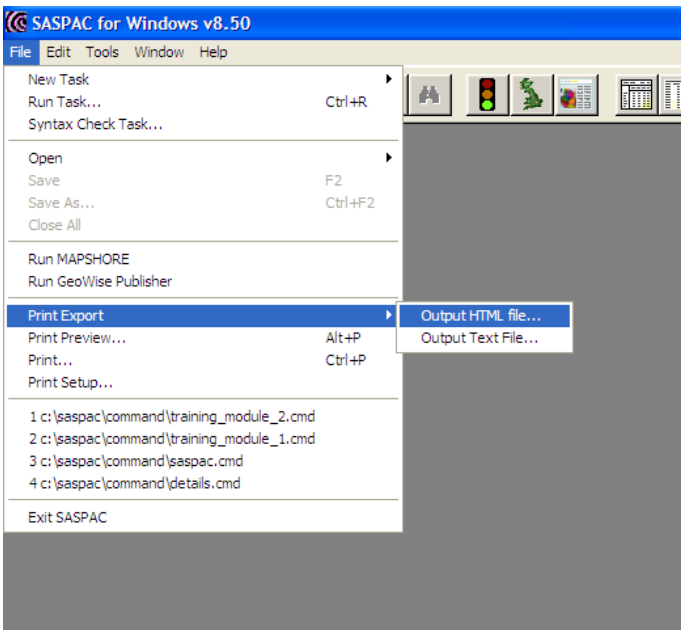
```
input system file name = "C:\SASPAC\Training /
System Files\2001 STANDARD TABLES FOR DISTRICTS /
IN LONDON.SYS"
input framework file = c:\saspac\frwork\st01.fwk
print tables ST052 ST103
output print file =
C:\SASPAC\REPORT\TRAININGMODULE2.PRN
end
finish
```

6. After the task has run the '**Print Preview**' Command Button may be used to examine the output before it is printed. This produces the output on screen as shown below.

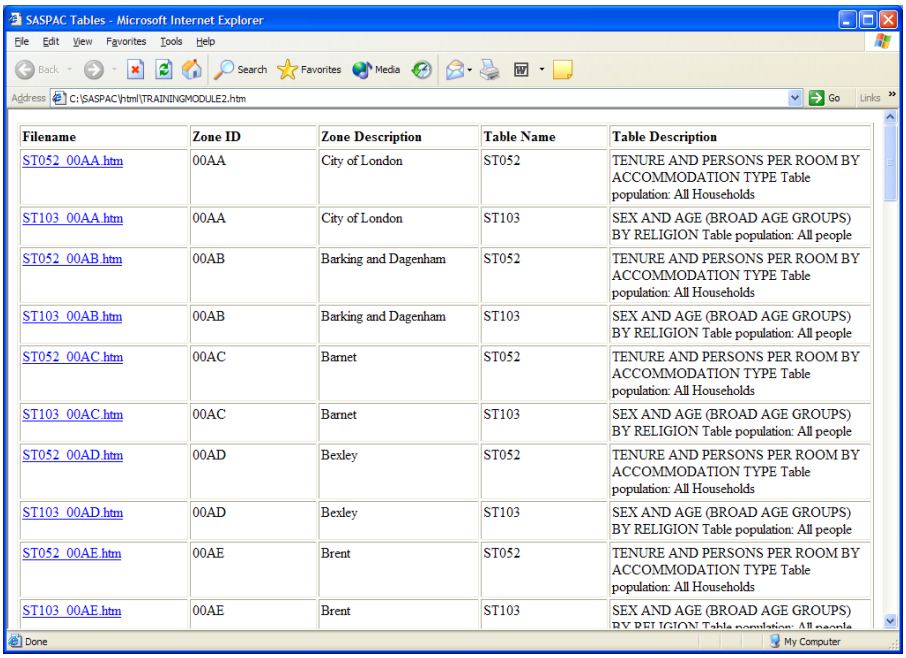


The banner heading on this page indicates that it is page 66, which happens to be the last page. This consists of two tables printed for each of 33 areas. This highlights the potential dangers in using SASPAC without the ability to filter out on areas. The current exercise has produced output for every area on the input system file. If that file had contained 1,000 areas which is entirely possible, then the output produced would have consisted of 2,000 pages. Selecting areas is thus a very useful tool within SASPAC and the next module deals with various ways of doing this.

7. The print file generated by this task can, if required, be converted to a HTML format by use of the '**File/Print Export/Output HTML**' file menu option as shown below.

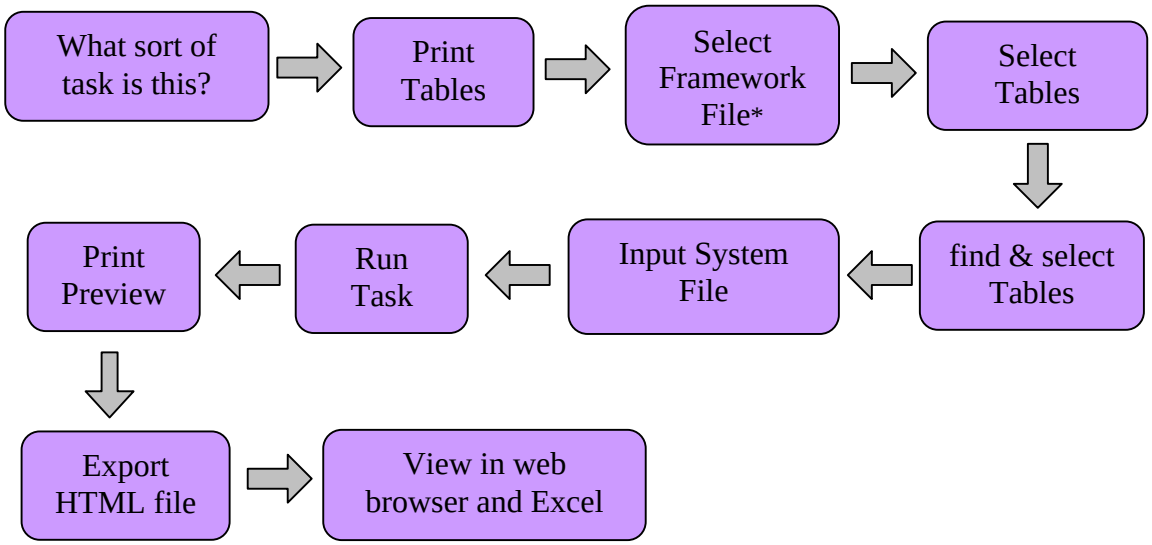


Once this route is followed, the user is presented with a set of Windows dialogue boxes which ask first for the name of the print file to be converted, and then the name and location of the index HTML file to be generated by this process. Once these files are located, the software will automatically create the necessary files. These consist of an index file and a set of files presenting each page of the print file as a separate HTML file. The index file contains hyperlinks to each of the page files as shown below.


SASPAC Tables - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Address C:\SASPAC\html\TRAININGMODULE2.htm
Filename | Zone ID | Zone Description | Table Name | Table Description
ST052_00AA.htm | 00AA | City of London | ST052 | TENURE AND PERSONS PER ROOM BY ACCOMMODATION TYPE Table population: All Households
ST103_00AA.htm | 00AA | City of London | ST103 | SEX AND AGE (BROAD AGE GROUPS) BY RELIGION Table population: All people
ST052_00AB.htm | 00AB | Barking and Dagenham | ST052 | TENURE AND PERSONS PER ROOM BY ACCOMMODATION TYPE Table population: All Households
ST103_00AB.htm | 00AB | Barking and Dagenham | ST103 | SEX AND AGE (BROAD AGE GROUPS) BY RELIGION Table population: All people
ST052_00AC.htm | 00AC | Barnet | ST052 | TENURE AND PERSONS PER ROOM BY ACCOMMODATION TYPE Table population: All Households
ST103_00AC.htm | 00AC | Barnet | ST103 | SEX AND AGE (BROAD AGE GROUPS) BY RELIGION Table population: All people
ST052_00AD.htm | 00AD | Bexley | ST052 | TENURE AND PERSONS PER ROOM BY ACCOMMODATION TYPE Table population: All Households
ST103_00AD.htm | 00AD | Bexley | ST103 | SEX AND AGE (BROAD AGE GROUPS) BY RELIGION Table population: All people
ST052_00AE.htm | 00AE | Brent | ST052 | TENURE AND PERSONS PER ROOM BY ACCOMMODATION TYPE Table population: All Households
ST103_00AE.htm | 00AE | Brent | ST103 | SEX AND AGE (BROAD AGE GROUPS)
Done
My Computer

8. Both the index file and the page files can be opened in either a web browser or in Excel. If a page file is opened in the latter, then the user can manipulate the table layout and data as necessary.

9. Summary of task sequence:

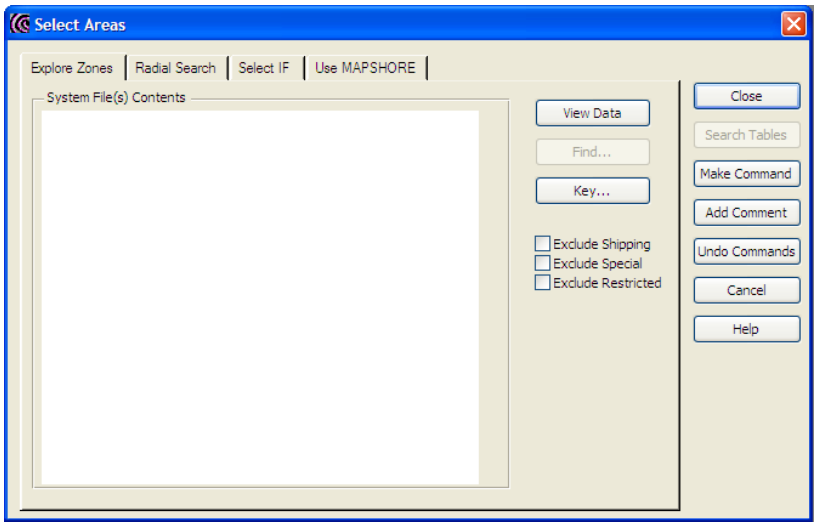


*make sure the system file, framework file and tables selected relate to the same dataset

Module 3A

Selection of Areas - by Text Listing

Whatever the task, the user will always have an option within SASPAC to select areas, normally achieved through use of the '**Select Areas**' Command Button on the right of the main task window. This will lead to a window as shown below, where there are four Tabs allowing selection of areas in different ways. This module covers use of the '**Explore Zones**' Tab which is the default, while the next module will cover the '**Use Mapshore**' Tab. Module 6 covers the two other options on this window.



Note that if a double geography file is input, this window will have slightly different tabs on display.

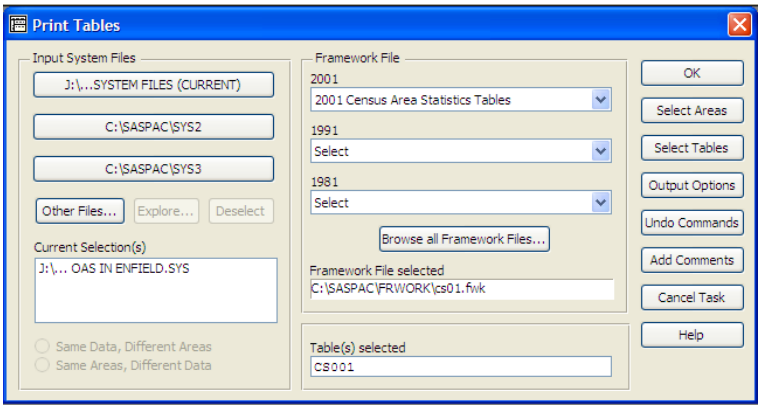
When the '**Explore Zones**' Tab is used, SASPAC will examine any input System Files to identify what areas are contained on them. For this functionality to work there must therefore be an input system file. This is one of the few cases where the order in which options are undertaken is important.

To access the list of areas found on file, the '**View Data**' Command Button is used. This initiates a procedure where the input system files are opened by SASPAC and their headers examined for a list of area codes. These are then presented to the users in the large text box which covers most of the window. If the areas on the file are hierarchical in structure, that is, they follow the Census Offices coding structure, the information will be presented to the user in a form which replicates Windows Explorer's manner of showing files within directories. Areas can then be selected for inclusion or exclusion.

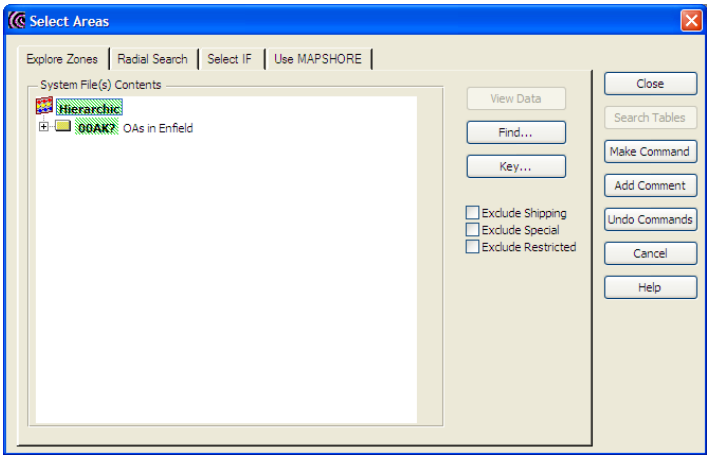
Note the three special categories of exclusion to the right of the text box. Also notice the '**Make Command**' Command Button. This button occurs in many places and must be used to convert selections into commands. If this button is not used, in this case, a prompting window is displayed which requests the user to make the command. However, there are instances where this is not the case, and in these circumstances, selections can be lost if the user neglects to click the '**Make Command**' Command Button.

Example: For the Output Areas of 00AKGP0004, 00AKGY0010 to 00AKGY0013, and 00AKHH0029 in the London Borough of Enfield, print Table CS001.

1. As in the last module, the task is defined as producing tables as printed output, and the same procedure is use for identifying the framework file, and the input system file. As we are asked to produce Table CS001 for OAs, we must input a Census Area Statistics file for OAs, and must also use the appropriate CS framework file. The table number has been specifically given, and so there is no need to enter the 'Select Tables' facility - the table number can be entered directly in the text box in the 'Print Tables' main window. All this is shown in the following screen image.



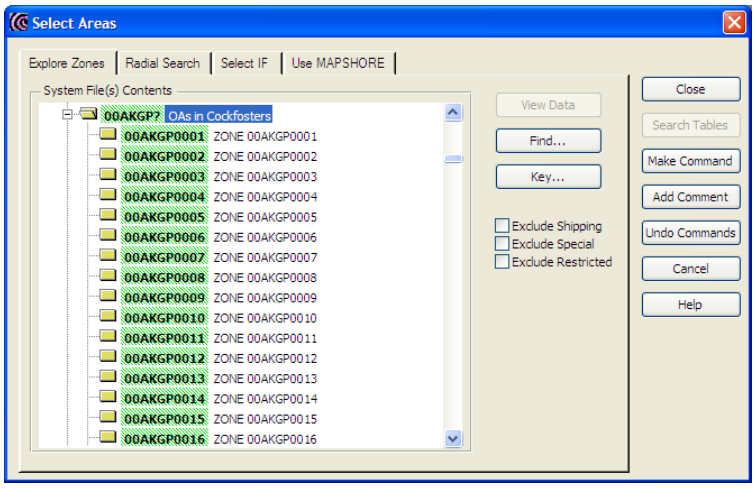
2. If the '**Select Areas**' Command Button is used, the window shown on the previous page is displayed to the user. When the '**View Data**' Command Button is now used, the software will open the input system file and display its contents as shown below.



This initial window shows that the software has found the highest level (or county) code of 00. As this is the code allocated to Unitary Authorities or Metropolitan Districts, the system has no information to identify this highest level. It must therefore look at the 3rd and 4th characters of the area code to obtain information on the area. In this case the four characters – 00AK – represent the London Borough of Enfield, and the codes consist of 10 characters. The system reports that it has found OAs in Enfield, as it has found

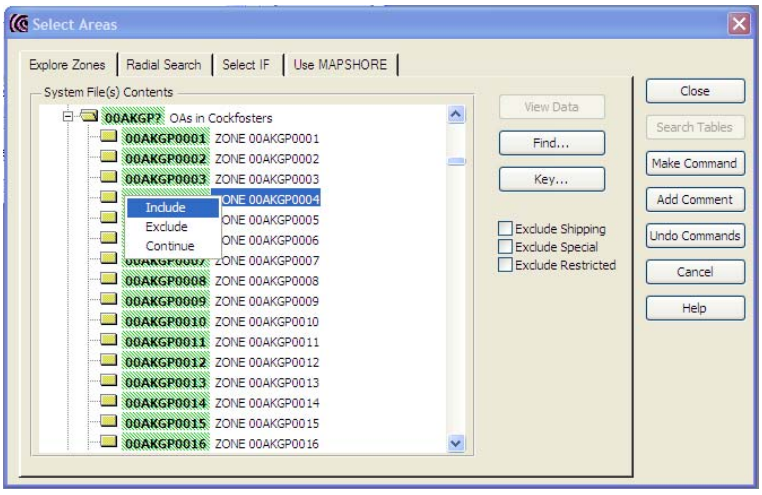
that the codes all have 10 characters. The '+' sign alongside the folder symbol indicates (as in Windows Explorer) that the folder can be expanded to display a lower level of area code.

This procedure can be continued until no further expansion is possible, and the codes of the individual areas found on the file are shown. The '?' symbol is used as a wild card, and if a code containing that symbol is selected then the software will automatically expand the code to the appropriate level. For example in this case, if the code '00AKGN?' is selected, then every area on the file that had a code which started with 00AKGN would be selected.



In this case, when the folders, or area levels, are expanded, the display in the window will be as shown in the next screen image. The areas have been expanded down through the hierarchy from county down to enumeration district. Sufficient information is now provided in the window for the required areas to be selected.

3. Using the scroll bar on the right of the text box, and closing unwanted are levels (or folders) allows the user control on what is displayed in the box, and allows access to the codes to be selected as shown in the next window. On highlighting the first code to be selected, a small pick list allowing the choice



of inclusion or exclusion is offered to the user. If 'continue' is selected, then the selection is ignored.

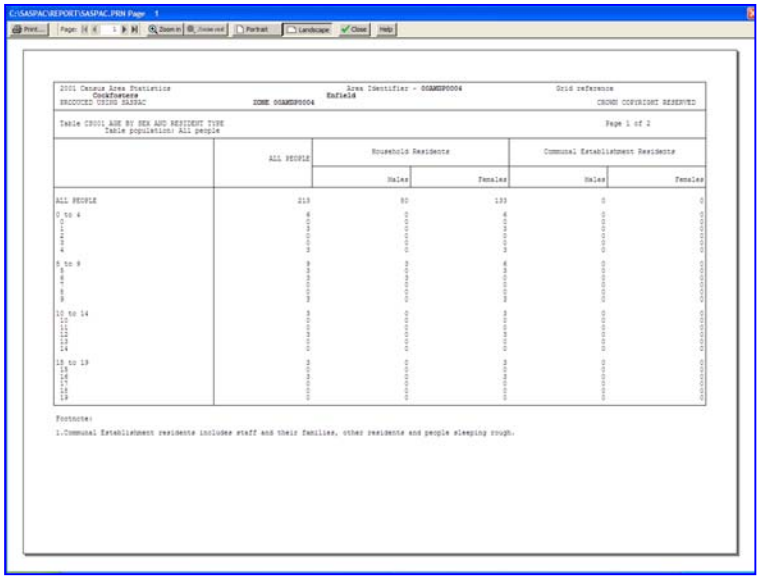
4. When the full list of required areas has been selected, the '**Make Command**' Command Button must be used to complete creation of the command. If the command is not created, the system will prompt the user to do so. After the '**Make Command**' Command Button is used, the 'Select Areas' window remains open to allow further selections to be made if necessary. The window is closed by use of the '**Close**' Command Button, at which point the user is returned to the 'Print Tables' main window, and the task may now be run.

5. If at any stage before final execution of the task, the user is unsure of what selections and choices have been made, then the '**Undo Commands**' Command Button, which is present on all windows, can be used to investigate the commands which have been generated up to that point. If the user finds that the wrong selection has been made then that command can be deleted. Alternatively, if a required selection has not been made, then the user can return to the option which would generate that selection or command.

6. The command file generated by the previous sequence is:

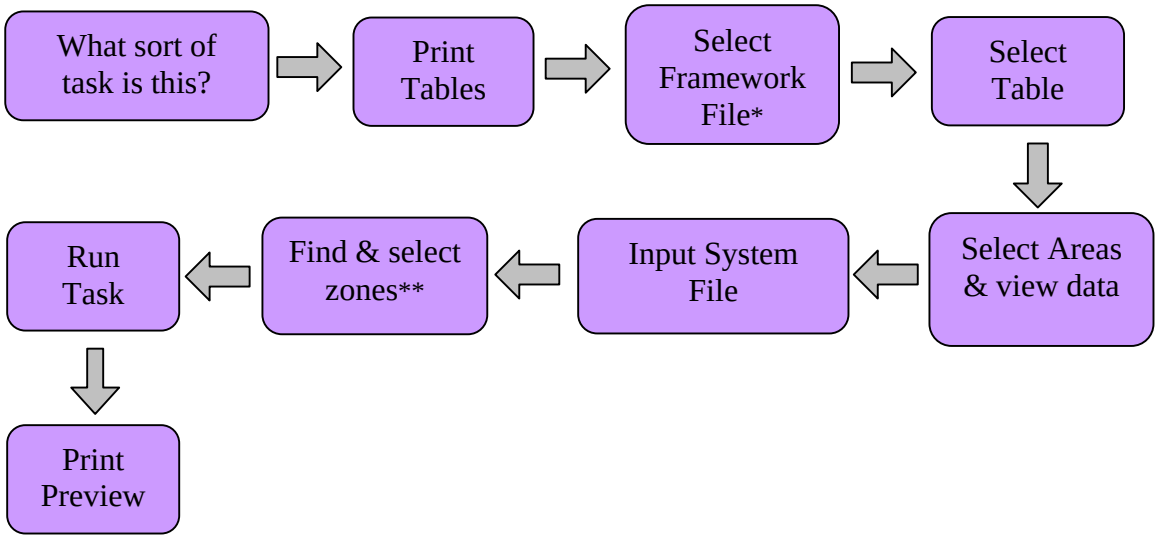
   input system file name = "C:\SASPAC\Training /
   System Files\2001 CENSUS AREA STATISTICS FOR OAS /
   IN ENFIELD.SYS"
   input framework file = c:\saspac\frwork\CA01.fwk
   include 00AKGP0004
   include 00AKGY0010 to 00AKGY0013
   include 00AKHH0029
   print tables CS001
   output print file = /
   C:\SASPAC\REPORT\TRAINING_MODULE3A.PRN
   end
   finish

7. After the task has run the '**Print Preview**' Command Button may be used to examine the output before it is printed. This produces the output on screen as shown below.



This command file only produces four pages of output - 1 table for each of four areas. By using area filtering or selection, the user is able to drastically reduce the quantity of output that is produced.

8. Summary of task sequence:



*make sure the system file, framework file and tables selected relate to the same dataset
**remember to 'make command' after your zone selections

Module 3B

Selection of Areas - Using Mapping

Version 5 of SASPAC for Windows added mapping capabilities to the existing functionality of the software following a link-up with Pebbleshore, the developers and owners of the Mapshore mapping package.

This additional capability allows the mapping element to be used in two ways:

As a front end browser allowing selection of areas from maps; and
As a generator of maps for presentation of results derived from SASPAC.

The ability to produce thematic maps and other graphical presentations of data is discussed in module 5, while this module covers the use of Mapshore in providing an easy-to-use method of identifying and selecting areas from on-screen maps thus obviating the need for large numbers of paper maps which over a period of years have usually become extremely dog-eared and tatty.

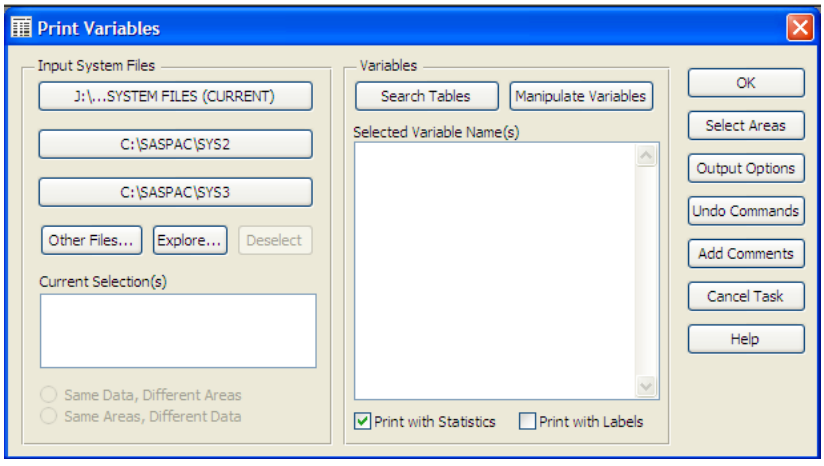
The base requirement for use of this facility is a file of digitised boundaries (at the appropriate level) in a valid format, and a raster map of the user's area as a bitmap file. Putting these two together gives users the full power of on-screen identification of areas.

There is no set location for these files as far as SASPAC is concerned, but the installation process creates a sub-directory of SASPAC named MAPPING which could be used for the process. The SASPAC.INI file does not set any defaults for mapping directories, and the MAPSHORE.INI file holds as default the last used directory.

When selecting areas using the mapping facility, the '**Use Mapshore**' Tab leads to a window with a '**Run Mapshore**' Command Button. Use of this button activates Mapshore in a separate window. The user then makes use of the facilities within Mapshore before transferring control - and the necessary information - back to SASPAC.

Example: For the area in the vicinity of the junction of the A105 and the A110 in Enfield, obtain the total population (CS0010001), and the total number of economically active residents (CS0210010)

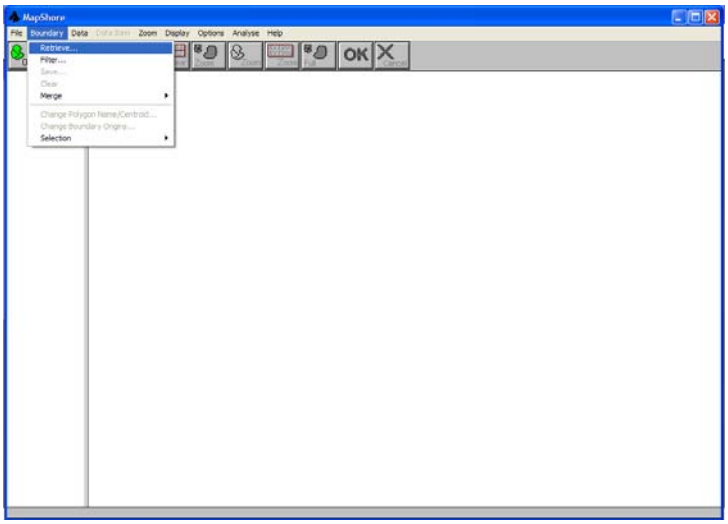
1. Again, the task is defined as producing variables as printed output, since the print variables procedure has the added parameter of allowing the production



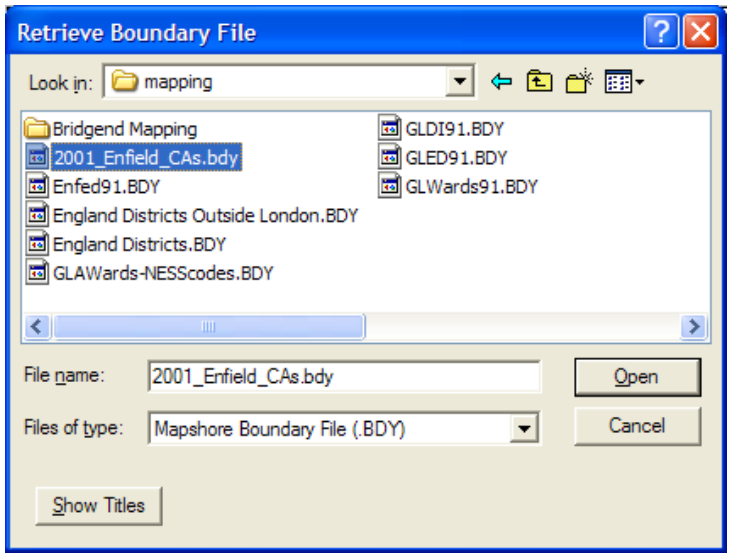
of statistics associated with the areas selected and printed. The task therefore becomes one of selecting all those areas lying partly or wholly within 0.5 kilometre of the selected point, and printing the required variables for these areas, using the '**With Statistics**' parameter. Selection of this parameter is as shown in the previous screen image.

2. When the '**Select Areas**' Command Button is then used, the 'Select Areas' window is presented at the default Tab ('**Explore Zones**'). To use Mapshore the '**Use Mapshore**' Tab must be used, and then the '**Run Mapshore**' Control Button is used to activate Mapshore.

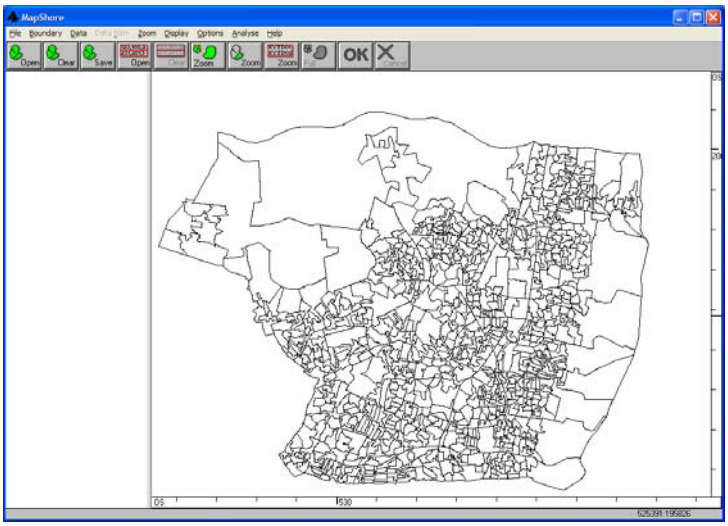
3. This action opens the Mapshore application and presents a blank window to the user. To start the mapping process, a boundary file must be selected. This is done through use of the **Boundary / Retrieve** menu as indicated in the next screen image.



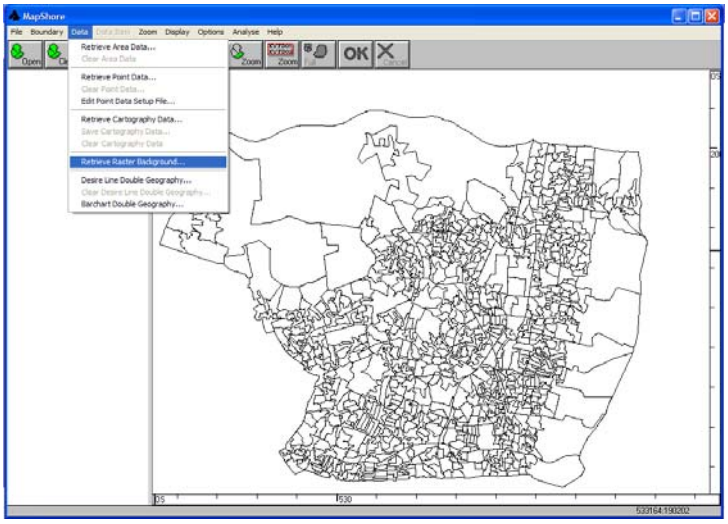
4. On use of this menu, a Windows Dialog box titled 'Retrieve Boundary File' is opened as shown below. The user then makes use of standard Windows dialog boxes to locate the required file, as shown below.



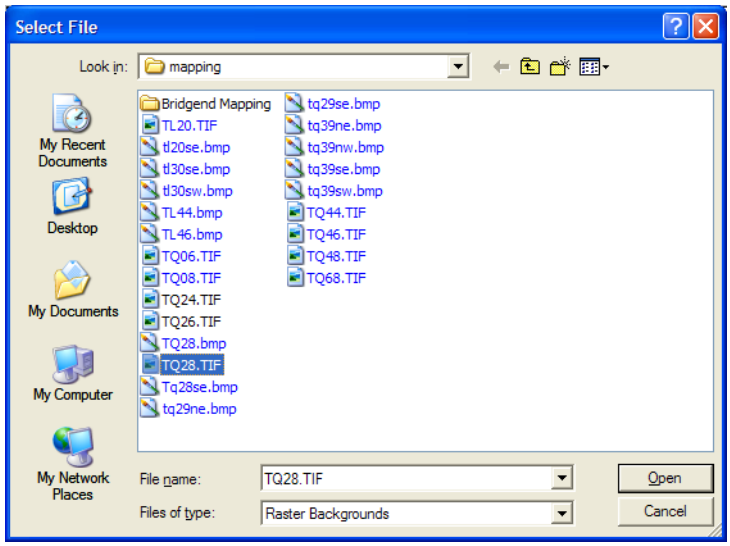
5. In this task, we are looking at small areas and so the boundary file we require is the Output Area file for Enfield (**2001_Enfield_CAs.bdy**), and opening this file will create a window as shown below. If files covering adjacent areas are required, then this may be done by repeating the above process, and the software will knit the two sets of boundaries together at the appropriate places.



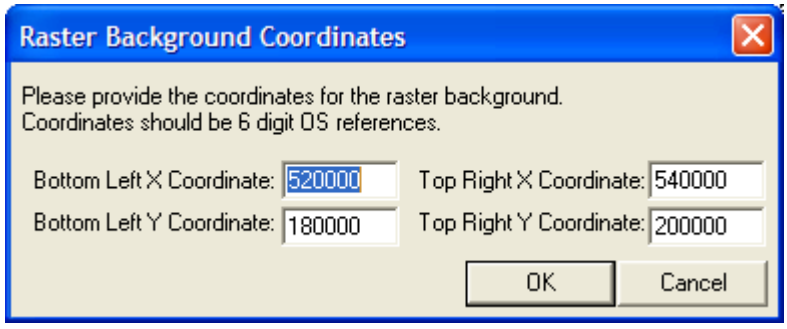
6. This image on its own is not enough, without some locational information which allows the user to identify precisely where the area of interest is located. This can be provided by a raster map selected through the '**Data / Retrieve Raster Background**' menu as shown here.



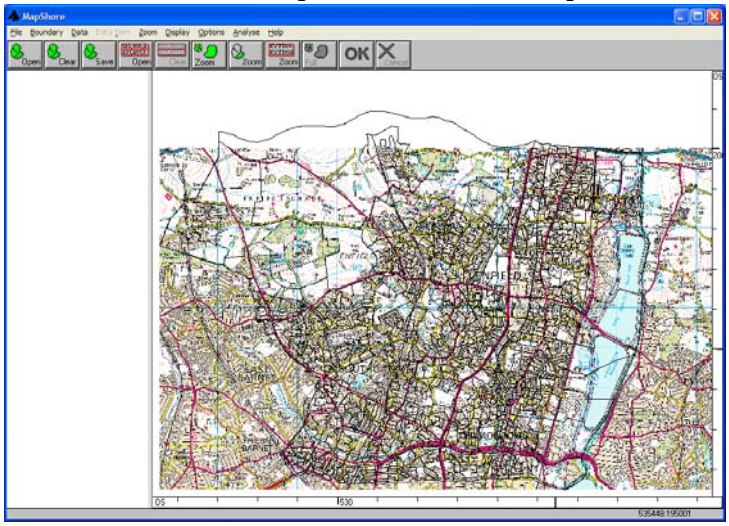
This menu option allows the selection of raster images as shown in the next screen image. Files of type 'Raster Backgrounds' are shown by the dialog box. Currently the raster types recognised by Mapshore are BMP and TIF. In this case the file selected is TQ28.TIF, and is an Ordnance Survey 1:50000 product. This is shown in the screen image on the following page.



7. In order to display the raster image in the correct position with reference to the boundary file, Mapshore must establish the coordinates of the bottom left and top right corners of the tile to be displayed. If the name of the file follows the Ordnance Survey's convention for naming 5km tiles - as in this case -, the software can deduce these points, as shown in the window bearing the title ' Raster Background Coordinates'. However, if these cannot be deduced from the file name, then the user must supply them in the appropriate text boxes.

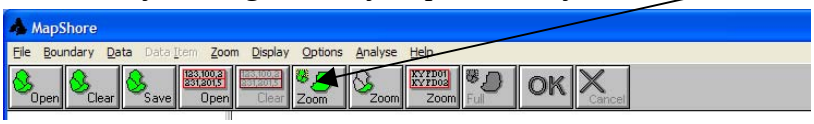


8. When this procedure is followed, the raster map will be placed in the correct position on the screen relative to the boundary file. Other raster files may now be opened, and these will also be placed in the correct position relative to the

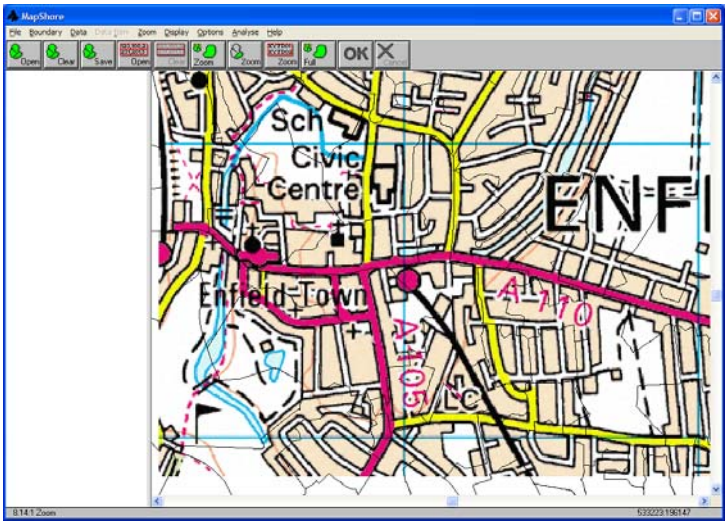


boundary file(s) and the existing raster map(s). The user may now zoom into a particular area of interest, or use the image shown to continue with the task.

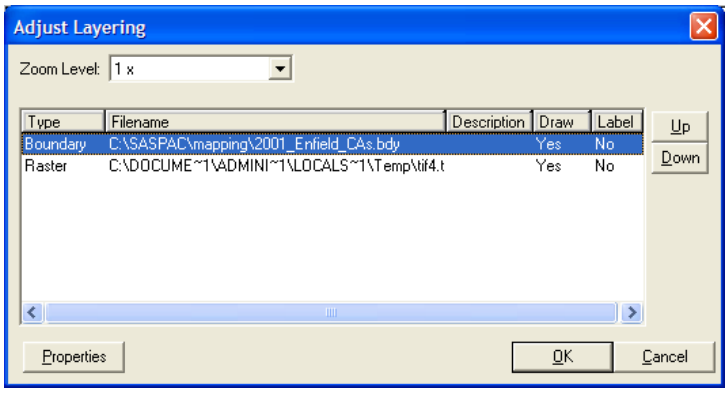
9. There is more than one-way of zooming within Mapshore, and in this case we will use the zoom by rectangle facility as provided by the sixth button at the



top of the screen. This allows the user to draw a rectangle covering the precise area of interest as shown on the following screen.



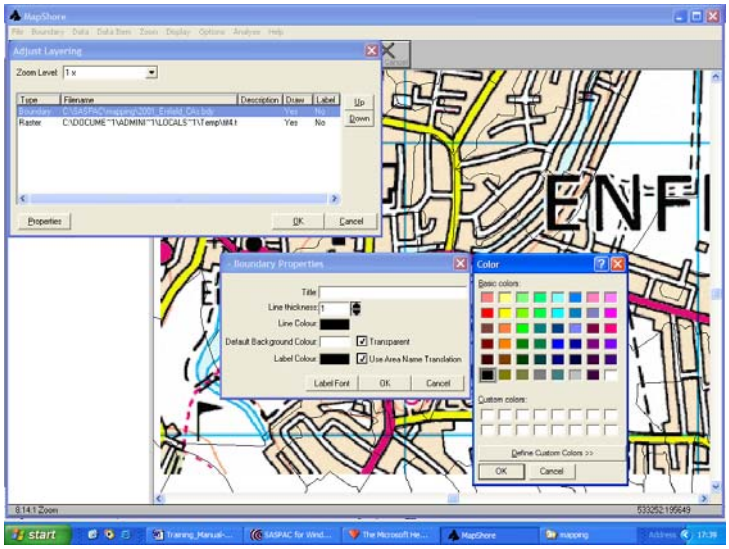
10 Once the necessary image is obtained, the Zoom / Layering menu option may be used to bring up the following window. In this window adjustments may be made to the image to achieve a better display.



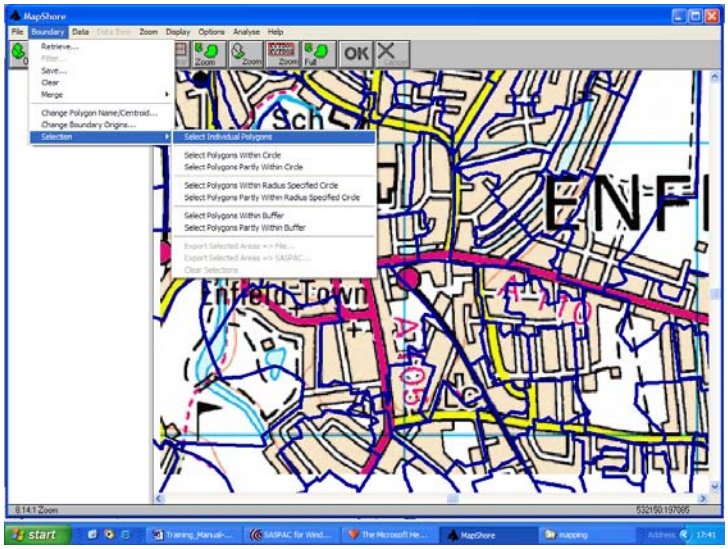
If necessary, the user can use this window to set the attributes of the open files at different levels of zoom. For example, the boundary lines could be set to a thickness of 1 at one zoom level, and at another level of zoom to a thickness of 3. Alternatively, the raster image could be invisible at the full size picture, but become visible after zooming in.

In this case, we will change the colour and thickness of the ED boundaries at this level. This is done by using the 'Properties' Command Button which allows access to a dialog box for changing some parameters and through this to a

colour palette for changing display colours. This is as shown in the next screen image.

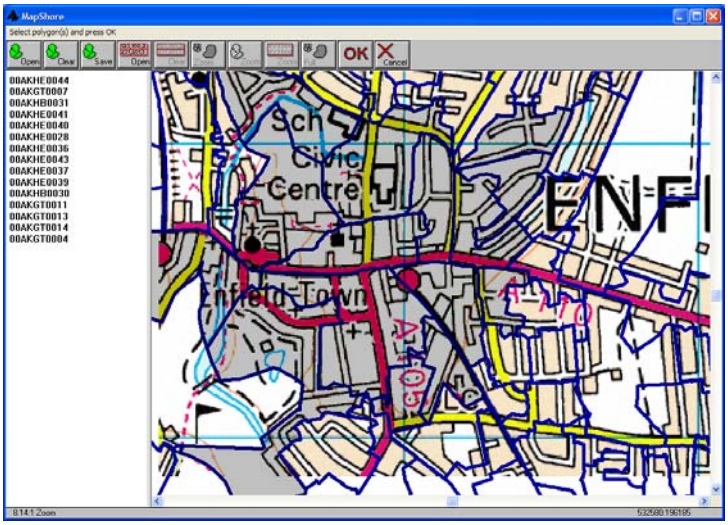


11. The result of this is shown below. This window also shows how the process of selecting areas for use within SASPAC is achieved. The **Boundary / Selection** menu item is used to access a menu defining various means of selecting areas - in this case we will choose the '**Select Individual Polygons**' item.

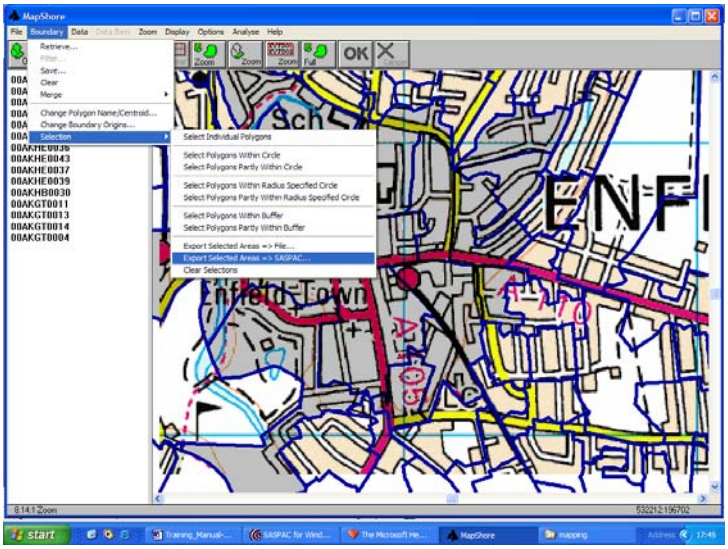


12. Having selected this option, the user must then point and click in the required areas (in this case, enumeration districts). The codes of these areas are then written to the text box on the left of the screen, and the selected areas are shaded on the map as shown on the screen image on the next page. Note that the selections may be toggled on and off. That is, clicking in an already selected area will deselect that area, both on the map, and in the text listing. If the user is satisfied with the selections made, then the '**OK**' button will be clicked - otherwise the '**Cancel**' button will be used. If the latter action is undertaken, then the selections are cleared and the user may use another selection procedure if necessary. However, if the selection is accepted, and the '**OK**' button is used, the system will await instructions as to what to do with the

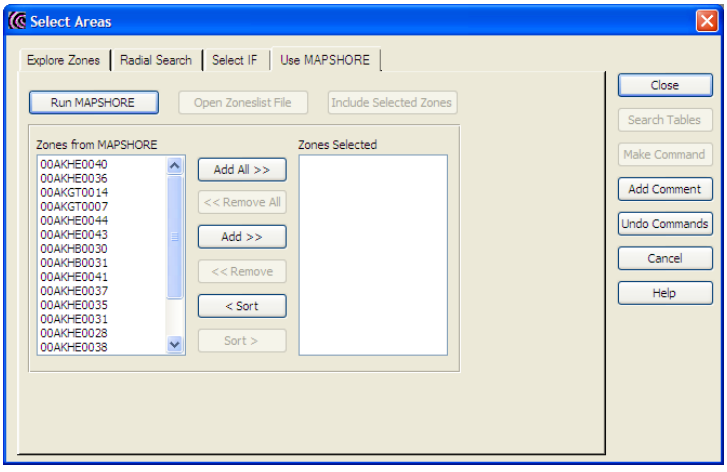
selection. In most cases, there will be a need, as in this instance, to transfer the information back to SASPAC.



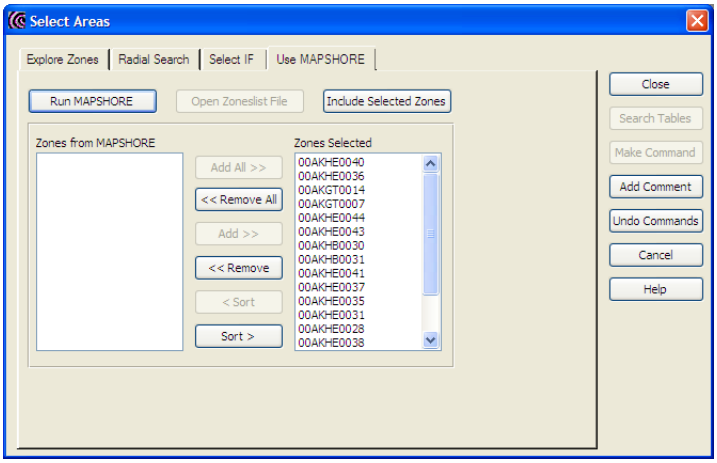
13. This is achieved by use of the 'Boundary / Selection menu, and selection of the 'Export Selected Areas => SASPAC…' item as shown below.



14. Mapshore's work is now complete, and control is transferred back to SASPAC at the point that it was left as shown below. The selected areas are now shown in the SASPAC window, and the user can use the selected areas as necessary.

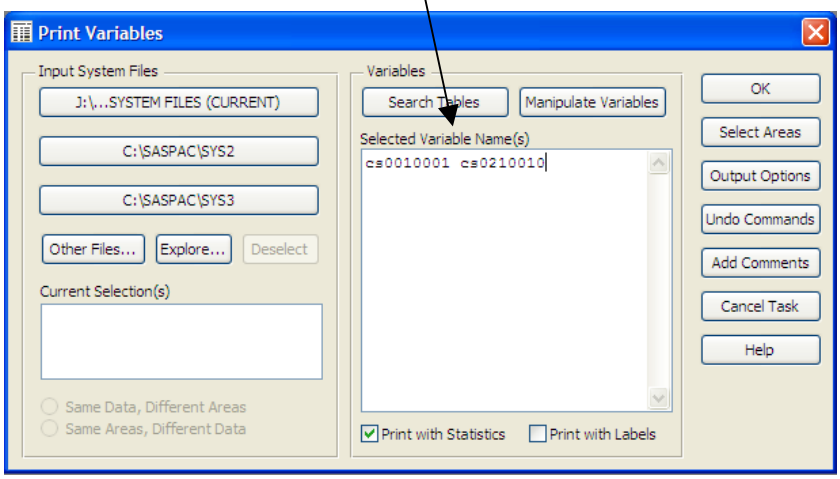


In most cases all the selected areas will be required, and so the '**Add All=>**' Command Button is used to transfer the area codes from the left box to the right box. If necessary, the '**Sort**' Command Button can be used to sort the codes into alphabetic order. The window below shows the result of using the '**Add All=>**' Command Button.



15. When the codes appear in the '**Zones Selected**' box, all that remains to be done is for the appropriate commands to be created. This is done by using the '**Include Selected Zones**' Command Button. Note that the '**Make Command**' Command Button will remain greyed out throughout this operation as the use of the 'Include Selected Zones' Command Button automatically creates the required commands.

16. Using the '**Close**' Command Button in this window will return to the user to the main task window. The final stage in this task is to ensure that the appropriate variable is output to the print file. If the variable identities are known, there is no need to use the '**Search Tables**' Command Button to find the identity, as they may be input directly to the '**Selected Variable Name(s)**' Text Box, as shown in this window.

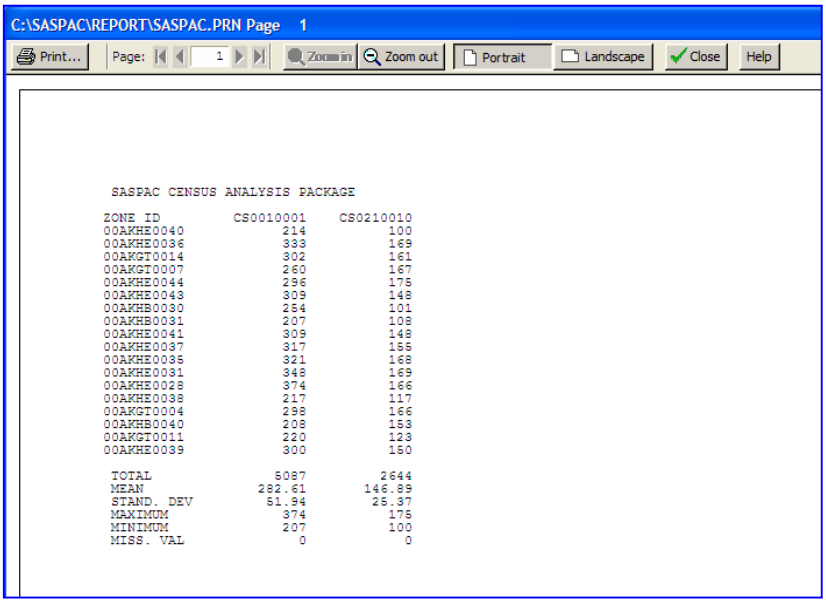


The task can now be run following the usual method of clicking on the '**OK**' Command Button in this window, and then on the '**OK**' command Button in the 'Task' window - having first allocated an user-defined name to the command file.

17. The command file generated by this sequence is as follows:

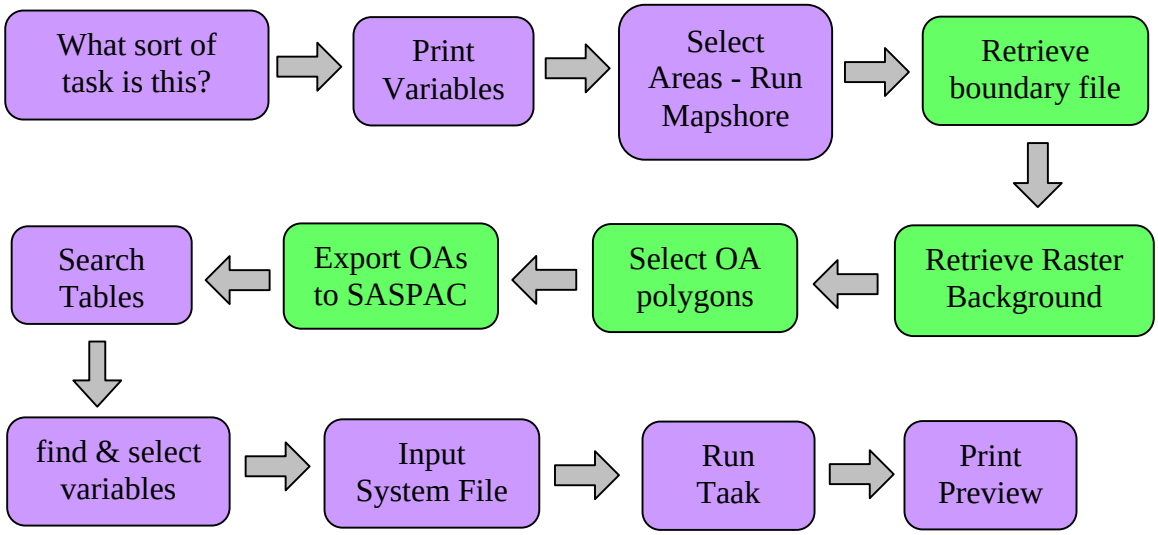
```
input system file name = "C:\SASPAC\Training /
System Files\2001 CENSUS AREA STATISTICS FOR OAS /
IN ENFIELD.SYS"
include 00AKHE0044
include 00AKGT0007
include 00AKHB0031
include 00AKHE0041
include 00AKHE0040
include 00AKHE0028
include 00AKHE0036
include 00AKHE0043
include 00AKHE0037
include 00AKHE0039
include 00AKHB0030
include 00AKGT0011
include 00AKGT0013
include 00AKGT0014
include 00AKGT0004
print variables with statistics CS0010001
CS0210010
output print file = /
C:\SASPAC\REPORT\TRAINING_MODULE3B.PRN
end
finish
```

and the output as viewed through use of the '**Print Preview**' option is shown below.



This form of the output where the '**With Statistics**' option was checked gives the total for the listed areas for each variable.

18. Summary of task sequence:



Module 4.

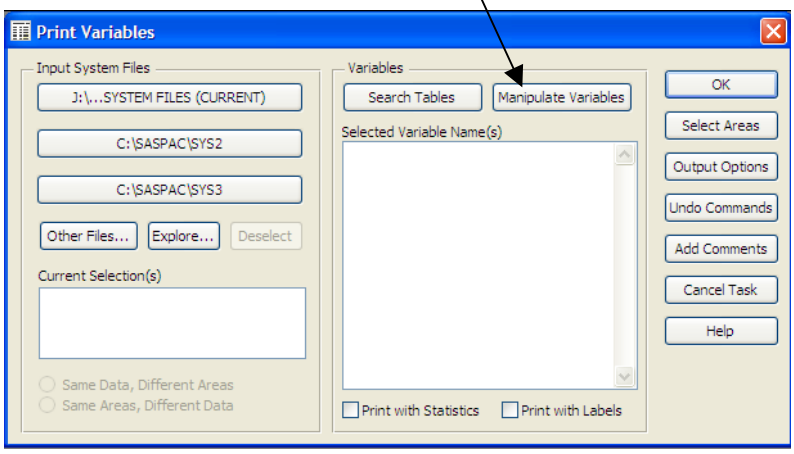
Creating New Variables

Even though the Census Area Statistics and Standard Tables between them contain some 75,000 variables, it is often the case that the information required is not directly available. The user may require a variable that is the sum of two or more other variables, or more commonly, may need to know a percentage or proportion. SASPAC provides facilities for these to be calculated and subsequently used in the same way as variables which already exist in the dataset.

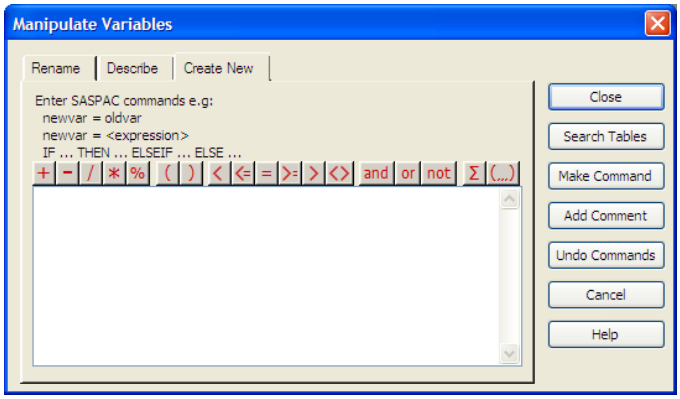
Calculations within SASPAC are undertaken using the standard mathematical operators such as '+', '-', '*', and '/'. The use of these operators is facilitated in the Windows system by a set of on-screen buttons.

Example: For all Output Areas in the London Borough of Enfield, calculate the percentage of the population that is aged 15 to 24, and output, along with the total population, to a print file. Ensure that the calculation is only done for areas where the total population is non-zero. Improve the column headings and change the report heading.

1. Again, the task is defined as producing variables as printed output, even though the variables required do not exist as such in the database. There is a need to create the variable which represents the required percentage. This is done through use of the '**Manipulate Variables**' Command Button which appears on the main task window.



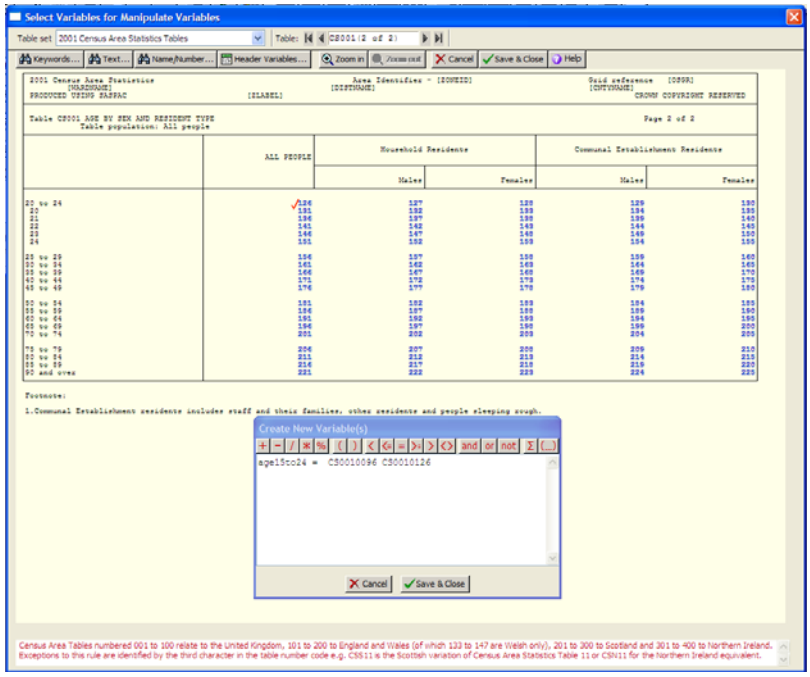
2. Use of this command button presents a window with three Tabs, and for this task, where we need to create a new variable, we need to use the third of these which bears the title '**Create New**' as shown in the screen image below.



3. In this window, the new variable is created by first allocating a descriptive name to it. This name can have a maximum of 12 characters which can only be numeric or alphabetic. No blanks or special characters may be used. We shall call it '**age15to24**', and this name is typed into the text box. We need to create a command of the form:

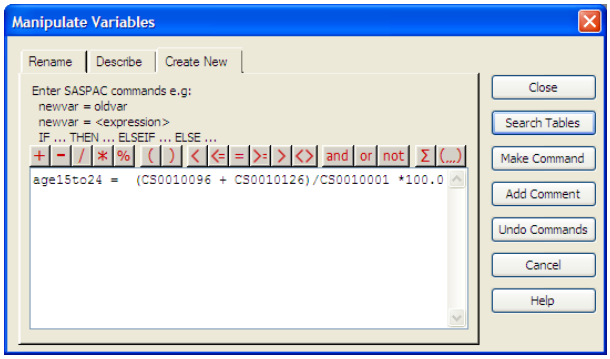
   age15to24 = <mathematical calculation>

   Assistance in creating this mathematical calculation can be provided through use of the '**Search Tables**' Command Button on the right of the window, and the symbol buttons above the text box. The Command Button performs the same functions as that to be found elsewhere, such as on the '**Print Variables**' main task window. Using it allows the user to select the required variables on-screen as shown below.

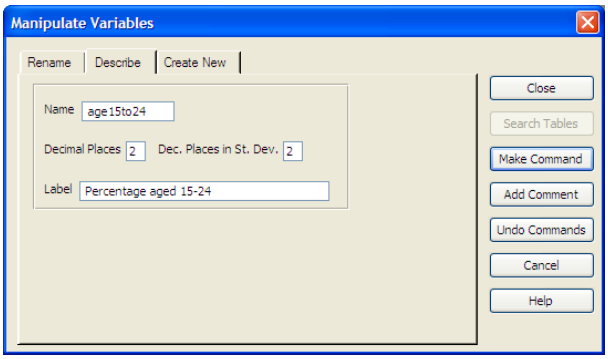


   Note that to create the total aged 15 to 24, we need the population aged 15 to 19 from page 1 of CS001, and that aged 20 to 24 from page 2 of the same table. To calculate the percentage, we then need the total population from page 1 of the table. By selecting the variables in this order, they are made available in the correct order for creation of the new variable.

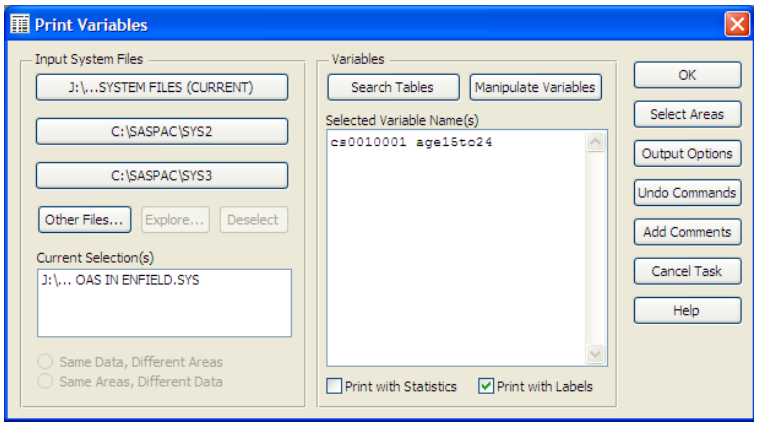
4. When the '**Save & Close**' Command Button is used, the selected variables are copied into the '**Manipulate Variables**' text box, where the appropriate mathematical symbols can be added as shown below.



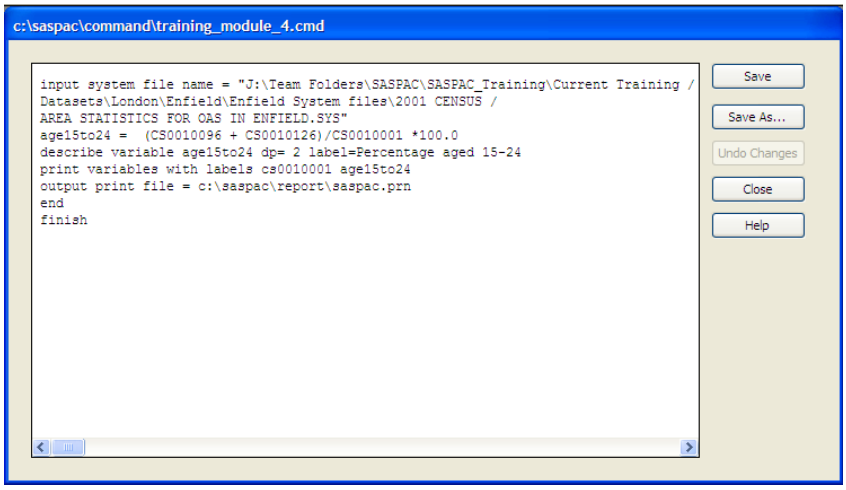
Once the command is complete, the '**Make Command**' Command Button must be used to create the command, otherwise the command will be lost. Once this button is pressed, the text box is cleared, and the user has the option to create another new variable. In this case we have created the only new variable that we need, and so we could now use the '**Close**' Command Button to return to the '**Print Variables**' main task window. However, one of the other tabs in this window allows us to describe a variable by adding a label. This is done as shown in the following window.



5. After making this command and closing the window, we must ensure that the appropriate variables are listed for output, and that a suitable input system file has been selected as shown in the image below. As we have described one of the variables, we can use the '**Print with Labels**' checkbox to make use of that label. The '**OK**' Command Button may be used to proceed to the 'Task' window.



6. In the 'Task' window, an user-defined name is allocated, but instead of using the 'Save and Run' Radio Button, we shall use the '**Save Only**' one, as there is a need to edit the generated command file to allow for the non-division by zero clause in the task. This is achieved by using the '**File / Open / Open Command File**' menu option to open a Windows dialog box in which the required file is identified and opened. (Alternatively, the file may be opened by accessing the '**Latest Files**' list under the '**File**' menu.). Whichever method is employed, the window shown on the next page will be presented to the user.



This is a text box into which any editing may be typed directly. Note the presence of two consecutive forward slashes '/' on the fourth line. One of these (the first) is the division symbol, while SASPAC has inserted the second as the line continuation symbol. Consecutive occurrences are not a problem, as SASPAC can differentiate between them. When SASPAC encounters a '/' symbol at the end of a line it will automatically add the text on the next line to the first line.

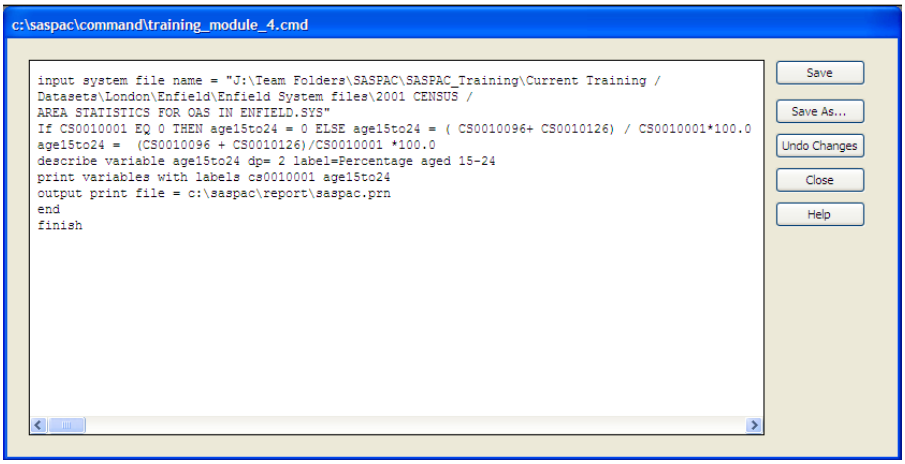
In this case we have the need to divide by CS0010001 only if the value of CS0010001 is not zero. This means that the line which reads:

age15to24 = (CS0010096+ CS0010126) / CS0010001*100.0

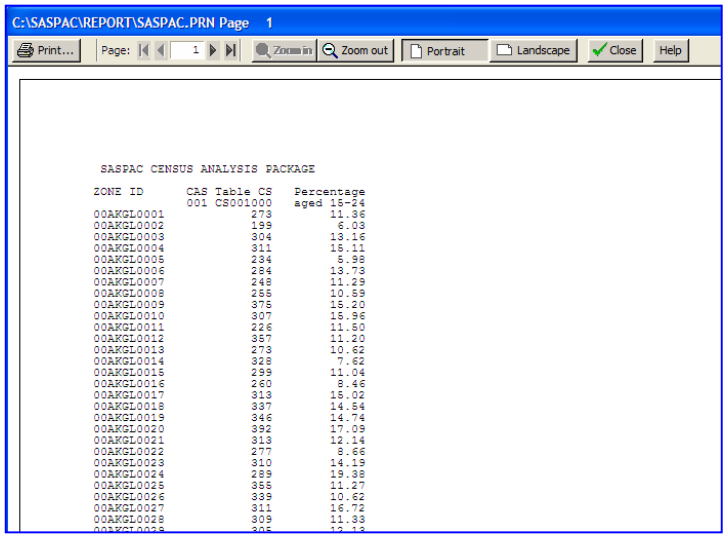
must be edited to read:

If CS0010001 EQ 0 THEN age15to24 = 0 ELSE
age15to24 = (CS0010096+ CS0010126) / CS0010001*100.0

as a single line as shown in the next image. The user need not have any concerns as to the length of the resulting line. SASPAC will split it with a continuation symbol if necessary.


c:\saspac\command\training_module_4.cmd
input system file name = "J:\Team Folders\SASPAC\SASPAC_Training\Current Training /
Datasets\London\Enfield\Enfield System files\2001 CENSUS /
AREA STATISTICS FOR OAS IN ENFIELD.SYS"
If CS0010001 EQ 0 THEN age15to24 = 0 ELSE age15to24 = (CS0010096+ CS0010126) / CS0010001*100.0
age15to24 = (CS0010096 + CS0010126)/CS0010001 *100.0
describe variable age15to24 dp= 2 label=Percentage aged 15-24
print variables with labels cs0010001 age15to24
output print file = c:\saspac\report\saspac.prn
end
finish
Save
Save As...
Undo Changes
Close
Help

7. When this command file is run by using the '**Save**' and '**Close File and Run Task**' menu option, the output that is obtained is as shown below. If the value of CS0010001 is zero for an area, the value of the new variable is set to 0.00, as defined, rather than the missing value symbol of '-'. Also note that the variable that was not described has the default label at the head of the column.

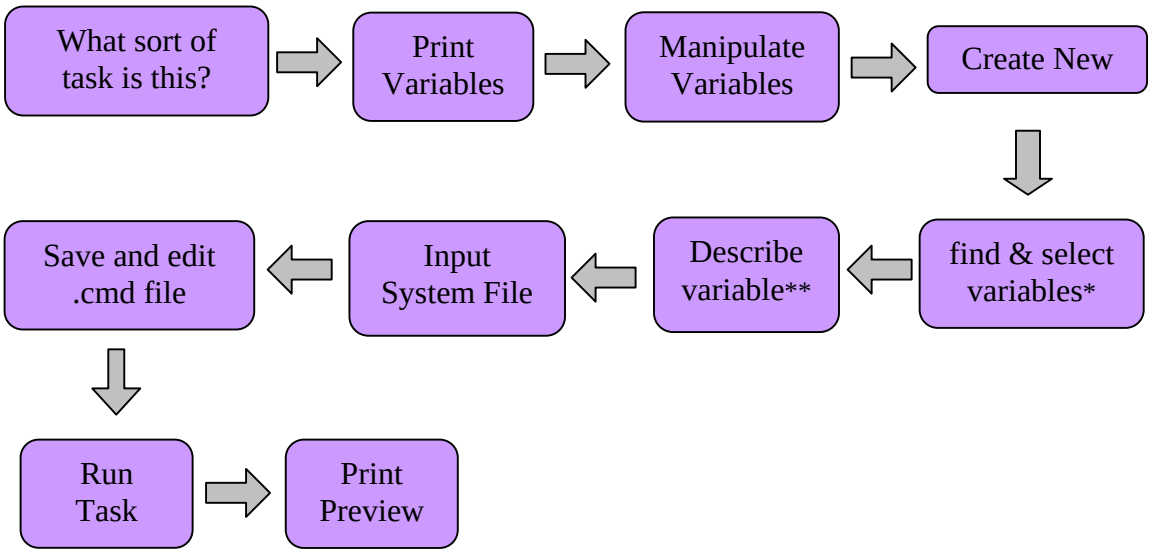


SASPAC CENSUS ANALYSIS PACKAGE

ZONE ID	CAS Table CS 001 CS0001000	Percentage aged 15-24
00AKGL0001	273	11.36
00AKGL0002	199	6.03
00AKGL0003	304	13.16
00AKGL0004	311	15.11
00AKGL0005	234	5.98
00AKGL0006	284	13.73
00AKGL0007	248	11.29
00AKGL0008	255	10.59
00AKGL0009	375	15.20
00AKGL0010	307	15.96
00AKGL0011	226	11.50
00AKGL0012	357	11.20
00AKGL0013	273	10.62
00AKGL0014	328	7.62
00AKGL0015	299	11.04
00AKGL0016	260	8.46
00AKGL0017	313	15.02
00AKGL0018	337	14.54
00AKGL0019	346	14.74
00AKGL0020	392	17.09
00AKGL0021	313	12.14
00AKGL0022	277	8.66
00AKGL0023	310	14.19
00AKGL0024	289	19.38
00AKGL0025	355	11.27
00AKGL0026	339	10.62
00AKGL0027	311	16.72
00AKGL0028	309	11.33

User Note.
There may be a need to edit the label by adding in blanks at suitable places to ensure that the label splits appropriately over two lines as 12 by 12 characters.

8. Summary of task sequence:



*remember to 'make command' after creating your new variable
**remember to 'make command' after describing your new variable

Module 5.

Output of Information from SASPAC

SASPAC's primary function is as a tool to provide quick and easy access to the very large datasets and allows small sub-sets of areas and variables to be identified for further manipulation and analysis. This further work may be undertaken within SASPAC or via some other item of software. As SASPAC was designed with limited analysis capabilities, the link to other analytic software is important.

In many cases, it will be sufficient for the user to obtain the results of the SASPAC interrogation as printed output, or as a simple text file. In such cases Report File output will be used. Alternatively, if the requirement is to save a sub-set of variables and/or areas for future access via SASPAC, then the output is saved as a System File. Finally, if the intention is to undertake further analysis in some other item of software, then the output will be in the form of an Interface File.

We have already discussed output of Report and System Files in previous modules. In this module we will discuss Interface Files and how they may be used to export data into other software for analysis and presentation purposes.

Interface files are SASPAC's method of allowing export of data to other software for further analysis. Generally, Interface Files will be output for:

- a selection of variables for all areas
- a selection of variables for a selection of areas.

Selection of areas and variables has already been discussed, and users should be aware that in addition to the standard count variables, the Header variables identified earlier may be also be output to Interface Files. Additionally, the ZONEID, which is automatically output to Print files and System files, is available for output, but is not automatically saved to Interface Files. Therefore, if there is a need for the ZONEID on the Interface File, it must be explicitly named on the SAVE VARIABLE command. If the command file is created through the Windows system then SASPAC will automatically prompt for the ZONEID if it is not present.

The interface file types available, and their descriptions are:

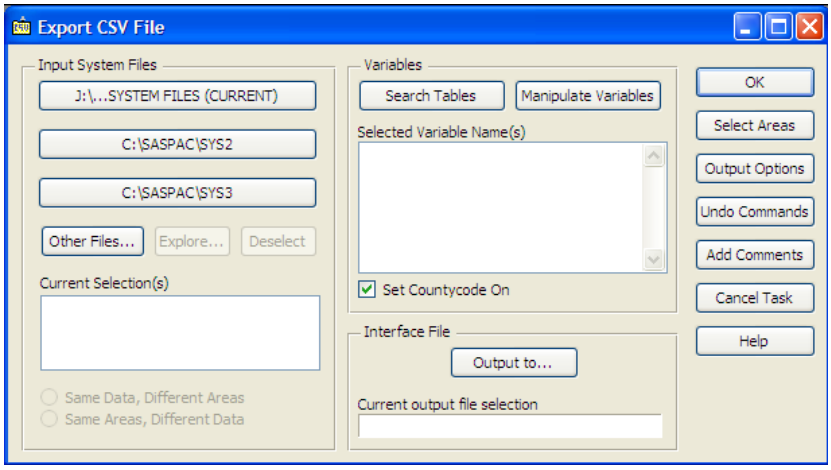
FORMATTED	An ASCII text file, where each output record has exactly the same format, with the format of the records being defined by a WIDTH command.
DELIMITED	An ASCII text file, where each variable is separated from its neighbour by a delimiter - normally a comma.
CSV	An ASCII text file, where each variable is separated from its neighbour by a comma.
WKS	A binary file suitable for export of data to spreadsheet software such as LOTUS 1-2-3, and Microsoft Excel.
DBF	A binary file suitable for export of data to database software such as dBase.
DIF	Data Interchange File suitable for importing data to a variety of software.
POSTCODE	A file, used in conjunction with a postcode directory, which allows output of estimated counts at postcode unit level.

When WKS, DBF, CSV, and DIF files are exported from SASPAC, the variable names are automatically exported along with the values. With Delimited, and Formatted files this must be specifically requested by the user.

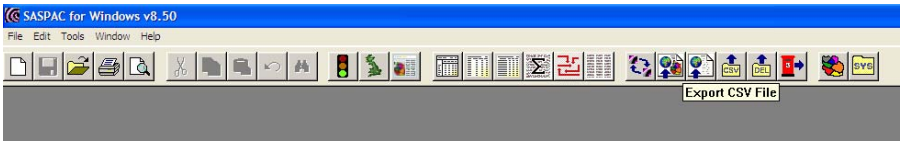
Two other types of Interface files are available - TRANSFER and MATRIX - but the use of these is restricted to specific instances, and they are not covered in this Module or in the SASPAC for Windows User Manual.

Example: Create a CSV file containing the percentage of the population aged less than 15 for all OAs in Enfield, and map the output. Export to Word as .jpg file and export in a .kmz format to view in Google Earth™.

1. The task here is defined as one involving the export of data to an interface file. The exported data can be mapped using SASPAC's own mapping utility - Mapshore. The initial selection route for this is '**File / New Task / Export Data / CSV File**'. Following this route will open the task window as shown below.



This window can also be reached through the Button Bar using the CSV button shown in the image below.



Note that in this window, the '**Set Countycode On**' checkbox is already ticked. This is because in most cases, the boundary files to which the data is to be mapped will contain the county code in the zone identifier. If the county code is not required this box can be unchecked.

2. This main task window is similar to others that have been used, and the facilities available through the Command Buttons are the same. In this case we wish to create a new variable to represent the percentage of the population that is aged 0 to 14. This is done through use of the '**Manipulate Variables**'

Command Button as in the previous module. The new variable is then calculated as:

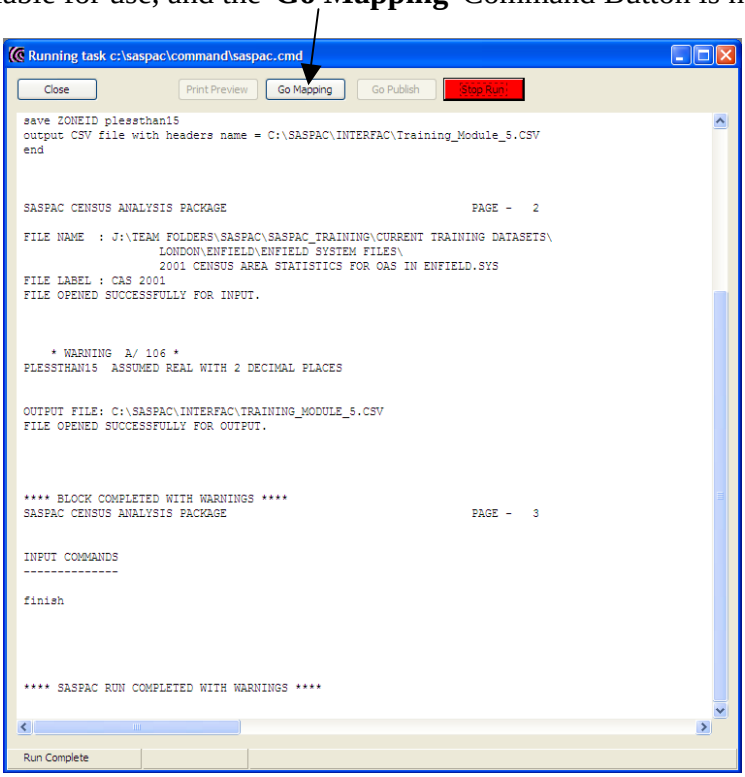
plessthan15 = (CS0010006+CS0010036+CS0010066)/CS0010001*100.0

When the new variable is defined, the names of the variables to be output are entered in the text box, and the necessary input system file is identified. The task may now be run, and the resulting 'Running Task' window is shown below.

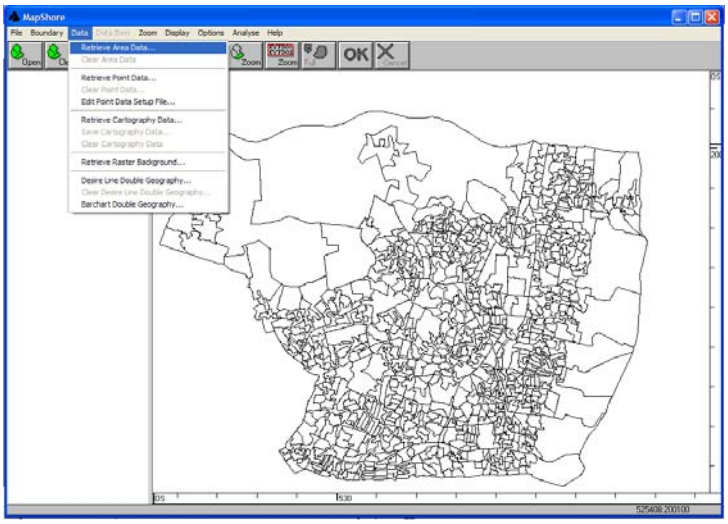
The Command File generated is as follows:

```
input system file name = "C:\SASPAC\Training /
System Files\2001 CENSUS AREA STATISTICS FOR OAS /
IN ENFIELD.SYS"
plessthan15=(CS0010006+CS0010036+CS0010066)/CS0010001 *100.0
set countycode on
save ZONEID plessthan15
output CSV file with headers name = /
C:\SASPAC\INTERFAC\Training_Module_5.CSV
end
finish
```

3. On completion of the task, the software detects that there is CSV file now available for use, and the '**Go Mapping**' Command Button is now made active.

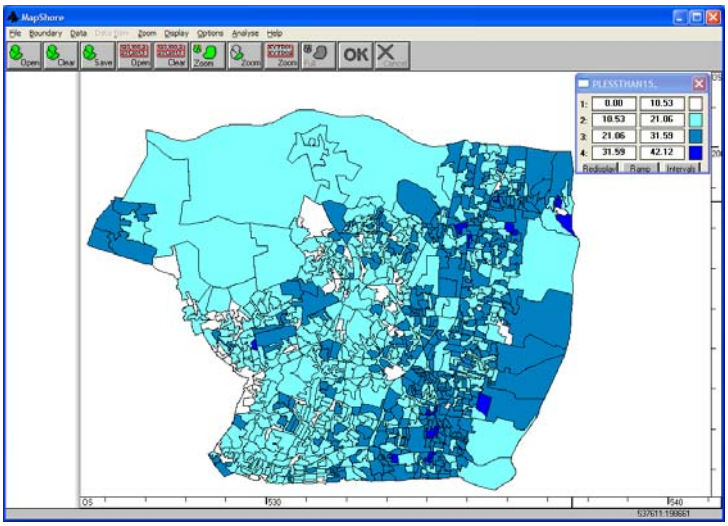

Running task c:\saspac\command\saspac.cmd
Close
Print Preview
Go Mapping
Go Publish
Stop Run
save ZONEID plessthan15
output CSV file with headers name = C:\SASPAC\INTERFAC\Training_Module_5.CSV
end
SASPAC CENSUS ANALYSIS PACKAGE PAGE - 2
FILE NAME : J:\TEAM FOLDERS\SASPAC\SASPAC_TRAINING\CURRENT TRAINING DATASETS\
LONDON\ENFIELD\ENFIELD SYSTEM FILES\
2001 CENSUS AREA STATISTICS FOR OAS IN ENFIELD.SYS
FILE LABEL : CAS 2001
FILE OPENED SUCCESSFULLY FOR INPUT.
* WARNING A/ 106 *
PLESSTHAN15 ASSUMED REAL WITH 2 DECIMAL PLACES
OUTPUT FILE: C:\SASPAC\INTERFAC\TRAINING_MODULE_5.CSV
FILE OPENED SUCCESSFULLY FOR OUTPUT.
**** BLOCK COMPLETED WITH WARNINGS ****
SASPAC CENSUS ANALYSIS PACKAGE PAGE - 3
INPUT COMMANDS
finish
**** SASPAC RUN COMPLETED WITH WARNINGS ****
Run Complete

4. Use of the '**Go Mapping**' Command Button opens the Mapshore application in the same way as was seen in Module 3B, and in the same way the first stage is then to open an appropriate boundary file - in this case the OAs for the London Borough of Enfield. When this is done, the CSV file may be opened using the '**Data / Retrieve Area Data**' menu item as shown below.



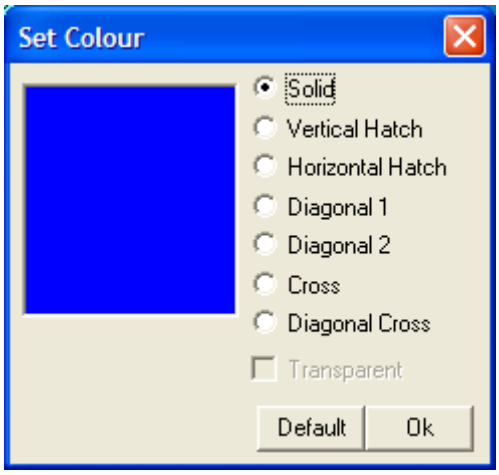
When this option is used, the user is asked to identify the appropriate data file for mapping. Once identified, if there is only one variable in the file, the thematic map will be drawn. If there is more than one variable in the file, then the user will be asked to identify the variable to be mapped.

The thematic map will be drawn using default shading and number of ranges. These defaults will be the values used the last time that Mapshore was used to draw a thematic map. The result in this instance is shown below.

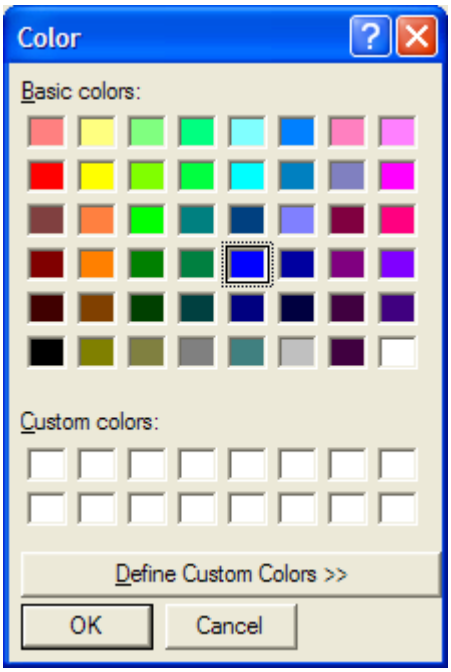


Once drawn the legend can be shown using the '**Display / Zonal Key**' menu item, and this can then be used to manipulate the displayed shadings and the number of ranges.

5. To change a colour, it is clicked in the legend and a '**Set Colour**' window is presented to the user as shown on the next page. In this window, the user can select the type of shading to apply to the map. If a non-solid shading is used, then the '**Transparent**' checkbox becomes active. If this is chosen, then any underlying image will become visible. To change the colour, then it is clicked

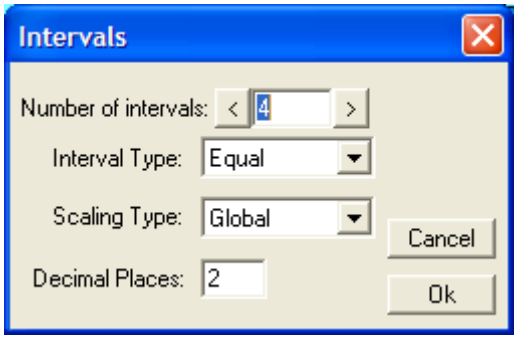


in this window, and a palette of colours is offered to the user, as shown in the next image.



A basic colour is selected, and if there is a need to refine this, then that can be done through use of the '**Define Custom Colors**>>'. This will then allow the user to create colours based on the mix of Hue, Saturation, and Luminosity, or as a unique combination of Red, Green, and Blue.

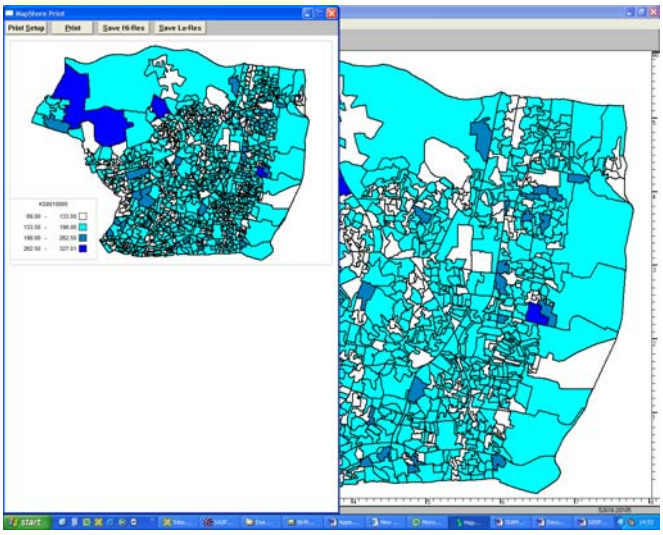
6. Numbers of ranges and type of ranges are set by using the appropriate buttons in the Zonal Key. If the '**Interval**' button is used, then the window shown here will appear. As can be seen, the Number of intervals, and the Interval Type are

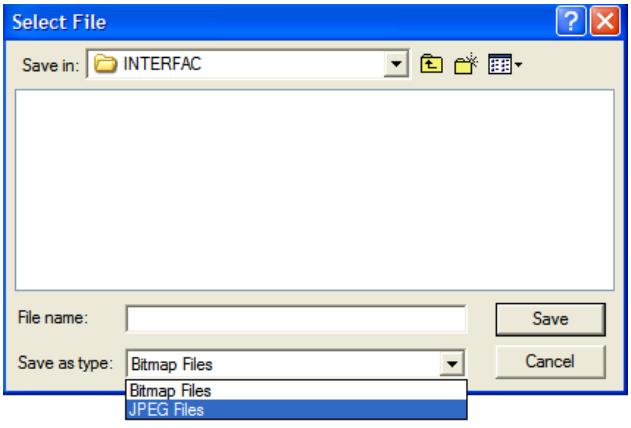


two of the parameters that may be set by the user. When the appropriate parameters have been chosen, the map is redrawn through use of the '**Redisplay**' button in the '**Zonal Key**' window.

7. Once a suitable result is obtained, the map may be printed, or exported using the '**File/Print...**' menu. This option allows a low or high resolution bitmap or jpeg image to be exported as shown in the two windows below. Chose to save a '**Hi-Res**' image in a jpg format with a suitable name.

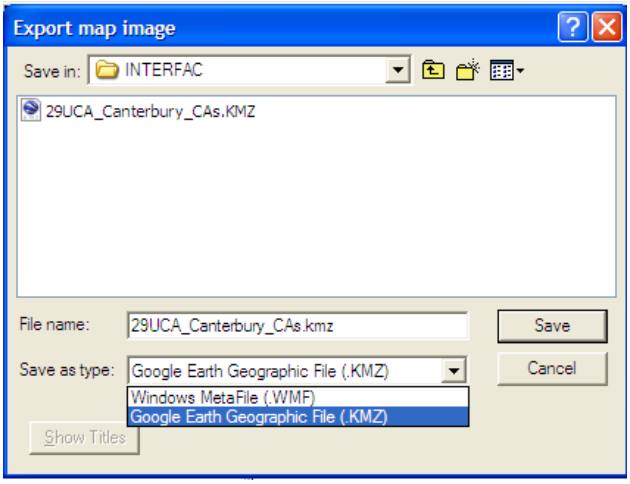
   This jpg file may then, for example, be pasted into a WORD document as the basis of a report or brief note related to the data depicted in the map. In WORD this would be done through use of the '**Insert/Picture/From File…**' menu item.





8. Alternatively, the release of SASPAC v8.5 saw the introduction of a Google Earth™ export facility. This tool allows the final Mapshore image to be exported in a .KMZ format, which can be viewed using Google's virtual globe browser 'Google™ Earth' (downloadable from: http://earth.google.com/).

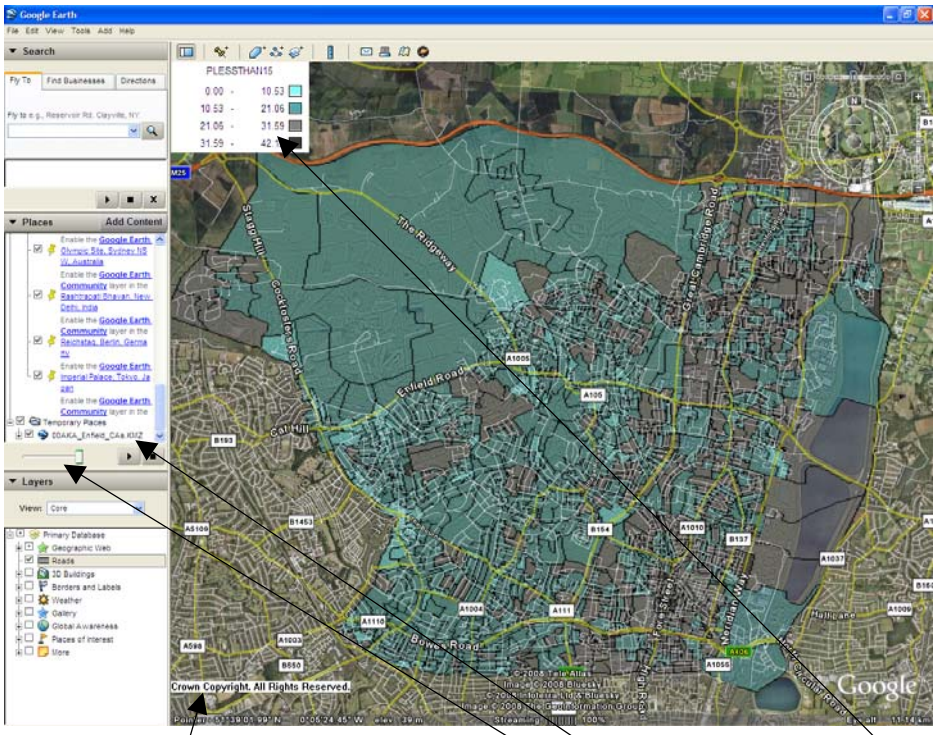
9. Use of the "**File|Export**" menu will open an export map image dialogue box as shown below:



10. Choose to save the file in a Google Earth Geographic File (.KMZ) format and give the file a suitable name. This opens a copyright statement window (shown below) in which the appropriate copyright statement for both the mapped Census data and the boundary file may be added. Finally, we must choose to save the file in a vector or raster format, in this case a vector format will be used.

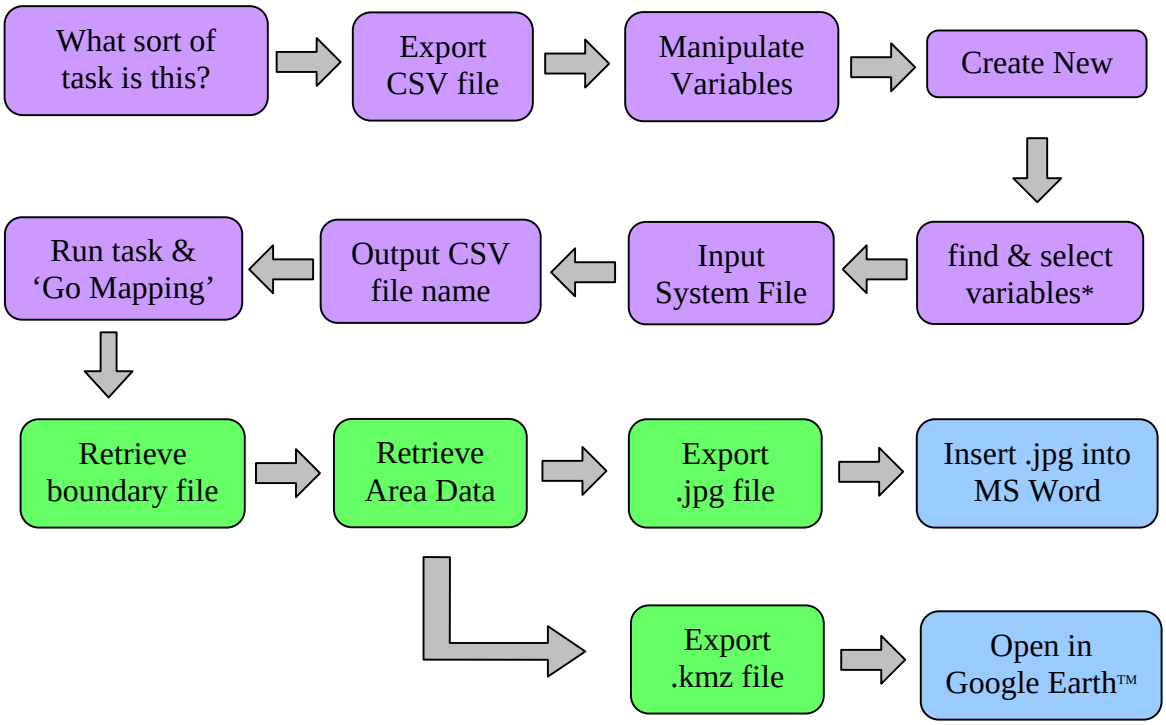
[Important: choosing to publish the .KMZ file in a vector format means other users will be able to extract and use the underlying geographic boundary data. It is important that you ensure this does not breach the terms and conditions of the licence agreement for these datasets. If you do not wish to make the geographic data available to 3rd parties, export the image in a raster format. If you are unsure of the licence restrictions, we recommend publishing in a raster format.]

11. Finally, the exported .KMZ file is opened by double clicking the file in Windows Explorer, using the left mouse button. This will launch the application and zoom to the position of the exported file on the Earth's surface. Alternatively, the "**File|Open**" menu within the Google Earth™ application itself can be used to open the .KMZ file directly. Whichever method is chosen, the result should look similar to the image below:



12. Here the final image has been placed in the correct geographical position of the LB of Enfield on the Earth's surface. Furthermore, the key and copyright statement are positioned in the top and bottom left corners of the map window, respectively. The menu to the left of the map window lists all the geographic features (OAs) individually. It is possible to display these, or not, using the tick-box next to the feature. Below this window is a slide bar that allows the transparency of the image to be adjusted from 0 to 100 percent. Finally, using the navigation bar in the top right-hand corner of the map window, it is possible to zoom in/out, rotate and adjust the tilt of the image. (To understand further some of the navigational aids and other datasets in Google Earth we recommend visiting: http://earth.google.com/tour.html).

13. Summary of task sequence



*remember to 'make command' after creating your new variable

Module 6.

Selection of Areas - SELECT IF

In Modules 4A and 4B, we discussed the selection of areas based on their area codes. In addition to selection based on the ZONEID of the required areas, SASPAC allows areas to be selected on the basis of criteria defined by the user and which apply to the value of selected variables. For example, selection may be restricted to all areas where the percentage of households living in local authority accommodation is greater than 50%.

The criteria for selection are constructed using the standard mathematical operators which are:

GT (or >)	Greater than
GE	Greater than or equal to
LT (or <)	Less than
LE	Less than or equal to
EQ (or =)	Equal to
NE	Not equal to

while complex conditions may be constructed using AND, OR, and NOT.

The criterion or criteria for selection is prefaced by the words SELECT IF, so that the format of the command is:

```
SELECT IF VAR1 GT 50
```

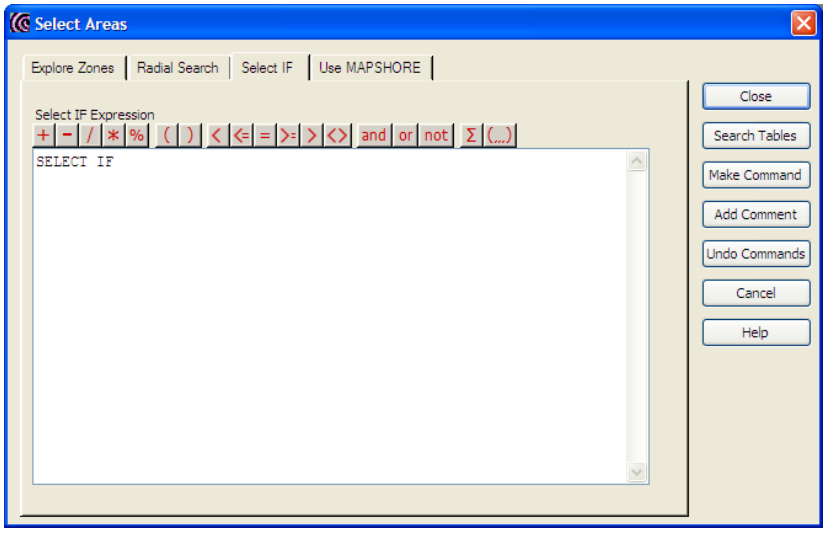
The functionality of radial searching is now available through the mapping environment.

Example: Output, to a print file, all OAs in the London Borough of Enfield, where the percentage of persons aged less than or equal to 15 is greater than or equal to 35%.

1. In this example, we are asked to create a new variable, and to output it to a print file only where a given criterion is satisfied. We must therefore define the task as one that outputs a report consisting of variables, and we must create a new variable to calculate the required percentage. The usual procedure is thus followed, where a main task window is defined by the '**File / New Task / Output Report / Print Variables**' menu option, or by use of the appropriate button on the Button Bar.

   The appropriate file for OAs in the London Borough of Enfield is opened, and the new variable is calculated using the '**Manipulate Variables**' Command Button, as described in Module 4.

2. The '**Select Areas**' Command Button is then used to access the '**Select Areas**' window which appears as the default on the '**Explore Zones**' Tab. Clicking on the '**Select If**' Tab will open the window shown here.



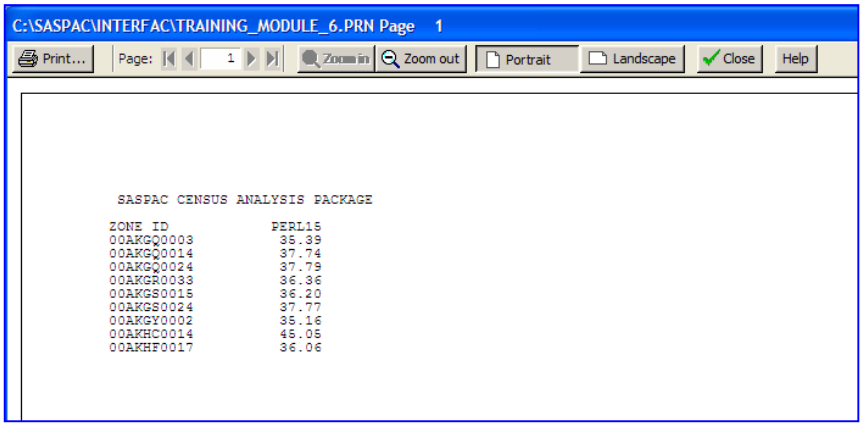
This window is very similar to the one which allows creation of new variables, and it may be used in a similar manner. However, it must be noted that user created variables cannot be selected from any list, and their identities must be typed into this text box. The arithmetic operators may be selected from the buttons above the text box, and combination criteria can be constructed using the logical operators.

3. When the text '**perl15 >= 35**' is entered into the text box, the '**Make Command**' Command Button is used to ensure the command is constructed and written to the command file. The user then returns to the main task window (assuming that no further constraints on output are required), where the identity of the required output variables is entered in the appropriate text box, or selected via the '**Search Tables**' Command Button. Once this is done the task may be run as before.

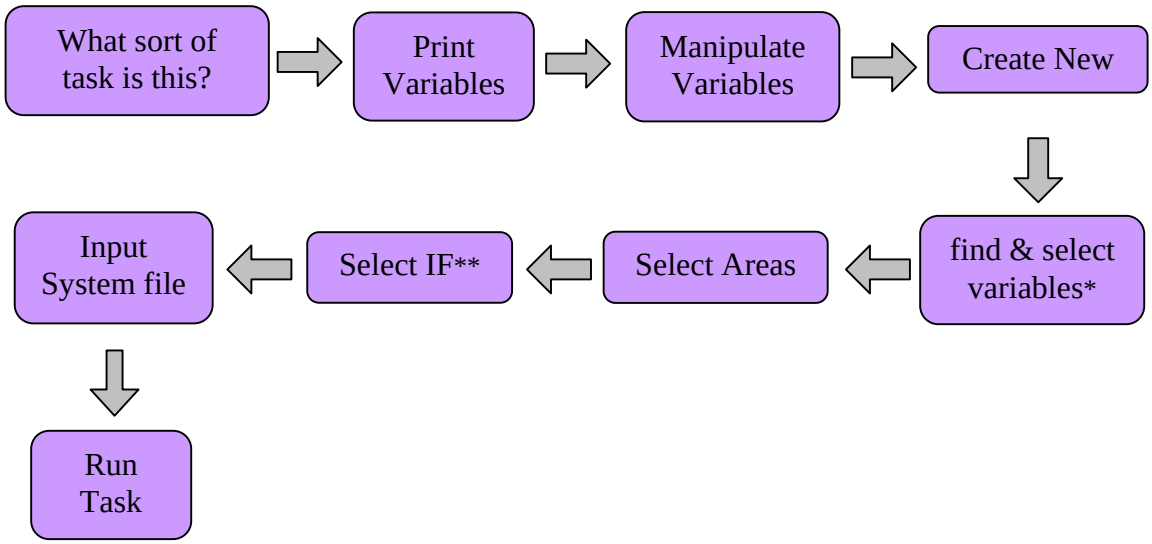
4. The command file generated by this sequence is:

```
input system file name = "C:\SASPAC\Training /
System Files\2001 CENSUS AREA STATISTICS FOR OAS
/
IN ENFIELD.SYS"
perl15=(CS0010006+ CS0010036+ CS0010066+/
CS0010101 ) / CS0010001*100.0
SELECT IF PERL15>=35
print variables perl15
output print file = /
C:\SASPAC\REPORT\TRAINING_MODULE_6.PRN
end
finish
```

5. The output from this task is shown below. By using the 'Select IF' clause, of the 800+ OAs in Enfield, only nine have been output as meeting the selection criterion.



6. Summary of task sequence:



*remember to 'make command' after creating your new variable
**remember to 'make command' after defining the Select IF criteria

Module 7A.

Creating New Zones - Individual Area Selection

As with variables, the areas of interest to the user may not be immediately available. However, SASPAC again provides the user with facilities for creating new areas from existing ones. This can be done in one of two ways:

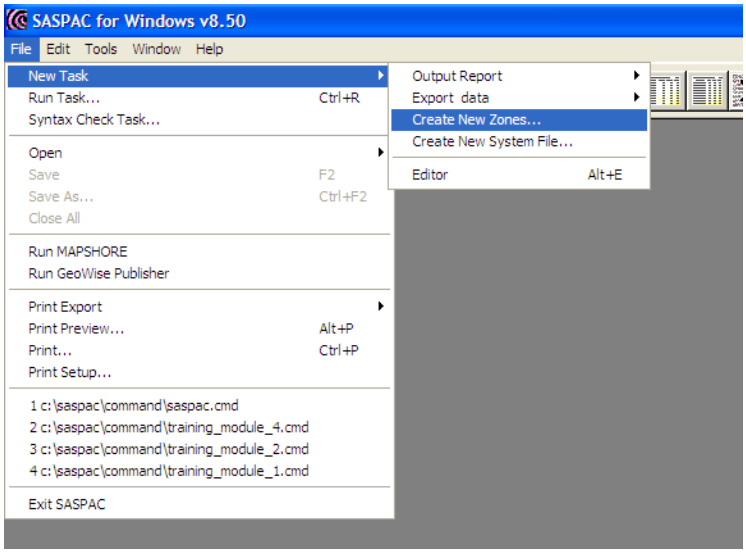
by combining explicitly identified existing zones; or
by reference to a gazetteer containing allocations of existing zones to new areas.

This Module deals with the former case, while the latter is dealt with by Module 7B.

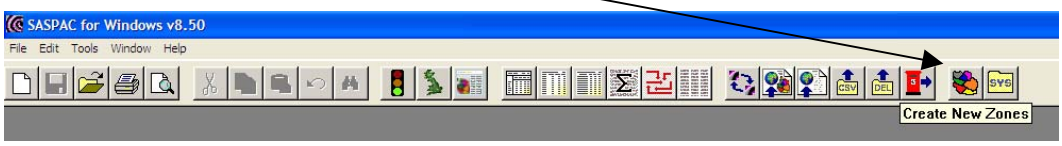
New zones may be created by adding together existing areas (which need not necessarily be at the same geographic level), or by subtracting an area from another area (the second area must be at a higher geographic level than the first). SASPAC also allows for scaling factors to be used to allocate proportions of existing areas to new zones. If a scaling factor is used, then it is applied equally to all the variables for the existing area.

Example: Create one new area as the aggregate of OAs 00AKGW0001 to 00AKGW0003, plus one third of 00AKGW0008, plus all of ward 00AKGT (Grange) less OA 00AKGT0007

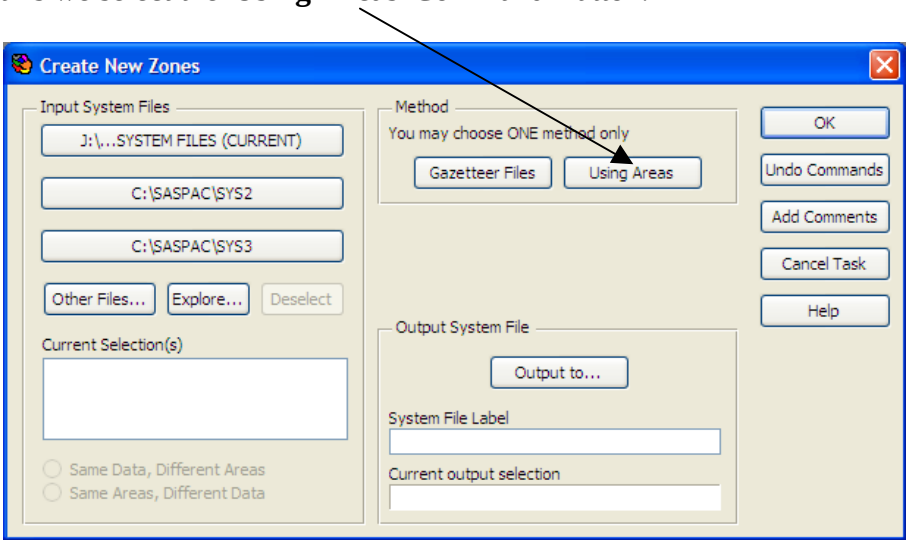
1. As the task here is to create new zones, it is by definition one of creating a new system file. When new zones are created, SASPAC must first save all the aggregated information to a system file, which is then available for interrogation in exactly the same way as the files containing standard areas. However, the menu system has a specific entry for creating new zones, and it is this rather than the creating new system file path that is followed. The menu path for starting this is shown in the following image.



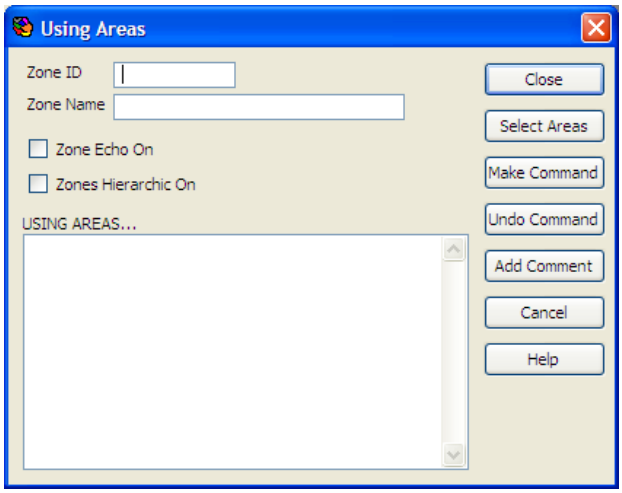
Alternatively, the '**Create New Zones**' button may be used as shown here,



2. Use of either of these options will open the '**Create New Zones**' main task window, as shown in the next image. This window is similar to those main task windows already seen in other modules. However, there is a major difference in the central area in that there are two Command Buttons which refer to the manner in which the new zones are to be created. In the current exercise we are concerned with creating new zones by explicitly identifying their codes, and to do this we select the '**Using Areas**' Command Button.



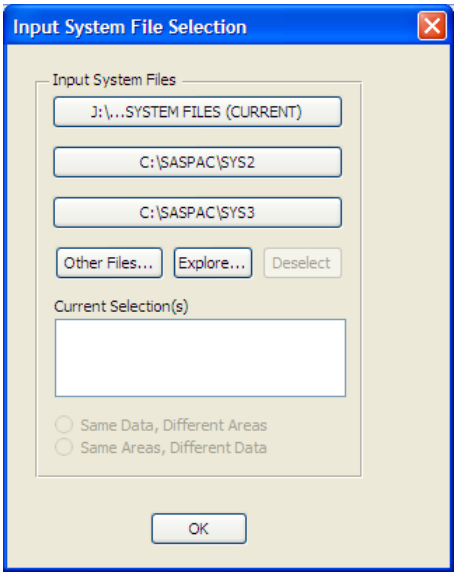
3. When this button is used the user is presented with a window (as shown below) which contains three text boxes into which the following information is to be entered:
   - the identity of the new area - up to 12 standard alpha/numeric characters without embedded blanks;
   - the name of the new area - up to 32 alpha/numeric characters including embedded blanks if required;
   - the identities of the existing areas which together will constitute the new area - the default operator is '+', so if there is a need for subtraction, that must be identified by a '-'.



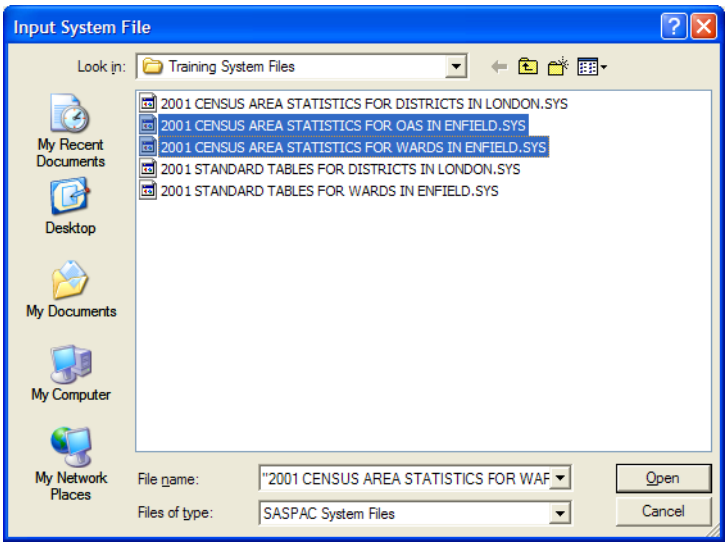
The first two boxes are completed by placing the cursor in the appropriate box and typing in the required information. The third box may be completed either by entering codes directly into the box, or through using the facilities discussed

in Modules 3A and 3B by use of the '**Select Areas**' Command Button. In this case we will use the text listing method, and so we must have selected input system files first.

If no system files have been specified when the '**View Data**' Command Button is used, then the software offers the user the opportunity to select those files through a modified window as shown here.

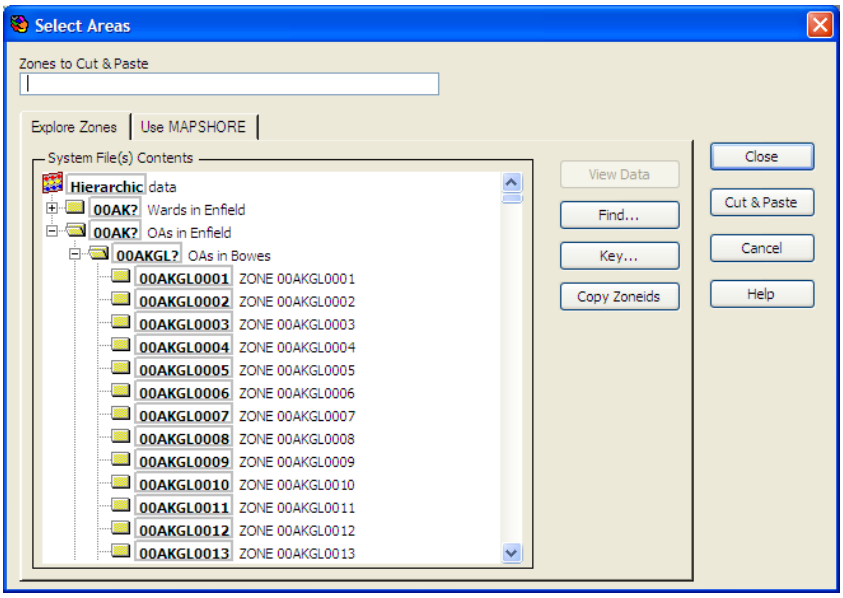


4. This window is effectively the left hand portion of the standard main task window, and the usual procedure is followed to select the file(s) and if necessary, the method of reading those files. In this case as we are combining wards and OAs we must input two system files, one for wards and one for OAs, as shown below.

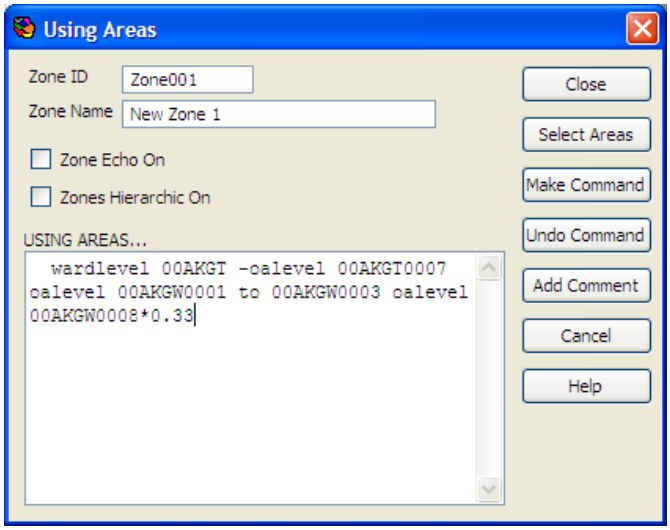


Note also that as we are combining the two area levels, we must use the lowest common denominator in terms of the datasets, which means that both have to be Census Area Statistics. We cannot mix the Census Area Statistics for OAs with the Standard Tables for wards.

5. As two input files have been selected, we must choose whether the files contain the same data, or the same areas, so that the appropriate '**READ IN**' command may be generated. In this case, it is the data that is common to the two files., and so the '**Same Data, Different Areas**' button is selected. When the input system files are selected, we can view the areas contained on them as shown here. The two files selected are shown to have OAs in Enfield and wards in Enfield on them. Use of the '+' symbols to open the hierarchy will display the OAs and wards in Enfield respectively.



The methods previously employed in Module 3A are now used to identify the required areas. When the full list is compiled, the '**Copy Selected Zoneids**' Command Button is used to transfer the highlighted identities to the text box at the top of the window, and finally the '**Cut & Paste**' Command Button is used to transfer these zone identities back to the '**Using Areas**' window. This empties the text window and allows the user to identify another set of identities if required. In this case we only want to create the single new zone, and so we use the '**Close**' Command Button to return to the '**Using Areas**' window. This window now contains the selected zone identities, and we can edit this text box to create the new zone as defined, and as shown below.

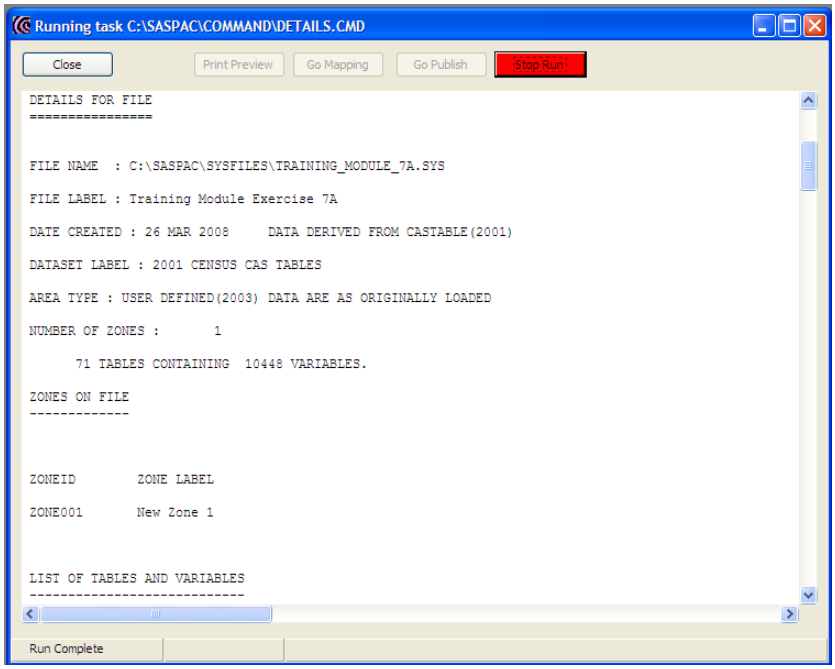


Finally the '**Make Command**' Command Button is used, followed by the '**Close**' Command Button to return to the main task window. At this stage, as we are creating a new system file, we must allocate it a name and a label, at which stage we can run the task. Note that the '**Task**' window only contains an option to rename the command file, as this time there is no output print file.

7. The command file generated by this sequence is:

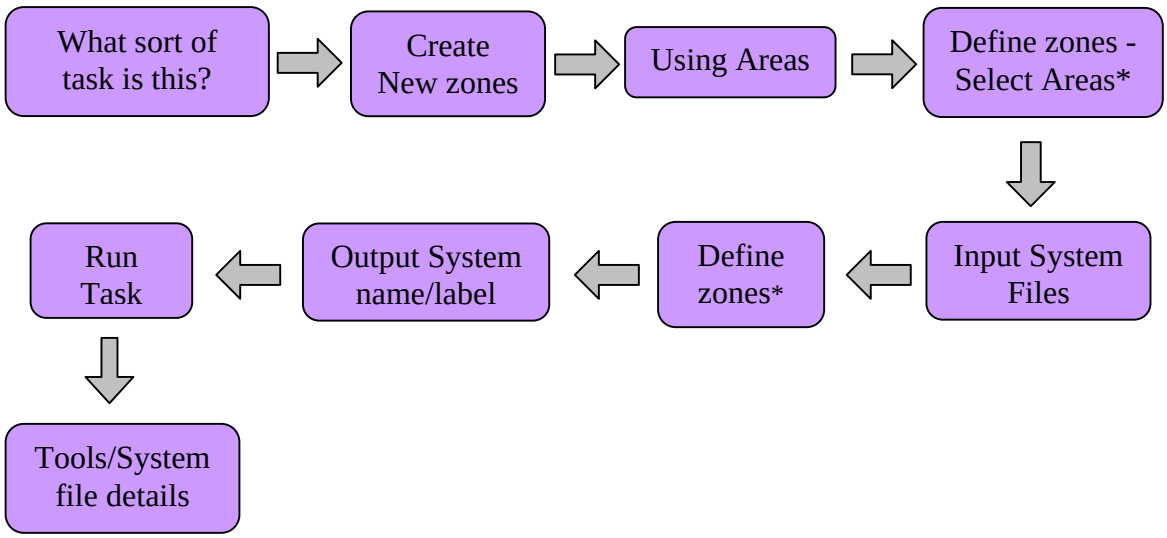
```
input system file name = "C:\SASPAC\Training /
System Files\2001 CENSUS AREA STATISTICS FOR /
WARDS IN ENFIELD.SYS"
input system file name = "C:\SASPAC\Training /
System Files\2001 CENSUS AREA STATISTICS FOR OAS /
IN ENFIELD.SYS"
read in series
new zone id = Zone001 name = New Zone 1
USING AREAS 00AKGW0001 to 00AKGW0003 /
00AKGW0008*0.33   -00AKGT0007   00AKGT
output system file name = /
C:\SASPAC\SYSFILES\Training_Module_7A.SYS /
label=Training Module Exercise 7A
end
finish
```

8. As the task only produces a system file output, there is no output that can be examined directly on screen. However, we can run tasks on this system file and we can use '**Tools / System File Details…**' to provide some information on the new files contents. If this is done, then the details of the file will be shown as on the following screen image.


Running task C:\SASPAC\COMMAND\DETAILS.CMD
Close
Print Preview
Go Mapping
Go Publish
Stop Run
DETAILS FOR FILE
FILE NAME : C:\SASPAC\SYSFILES\TRAINING_MODULE_7A.SYS
FILE LABEL : Training Module Exercise 7A
DATE CREATED : 26 MAR 2008 DATA DERIVED FROM CASTABLE(2001)
DATASET LABEL : 2001 CENSUS CAS TABLES
AREA TYPE : USER DEFINED(2003) DATA ARE AS ORIGINALLY LOADED
NUMBER OF ZONES : 1
71 TABLES CONTAINING 10448 VARIABLES.
ZONES ON FILE
ZONEID ZONE LABEL
ZONE001 New Zone 1
LIST OF TABLES AND VARIABLES
Run Complete

This screen image shows the result of such a check, and it shows that the file created contains the single zone, and 71 tables, as expected.

9. A summary of the task sequence:



*remember to ‘make command’

Module 7B.

Creating New Zones - Gazetteer Files

Sometimes, when new zones are to be created, their definition in terms of existing areas may already be available in a machine-readable format as a constitution file. Such a file may be introduced to SASPAC as a Gazetteer File - subject possibly to some editing.

A Gazetteer File is an ASCII text file in a user-defined format where each line of the file links an existing area (or part of an existing area) to a new area, through their zone identities. A typical Gazetteer File might look like the following:

```
00AKGW0001   AREA01
00AKGW0002   AREA02   0.25
00AKGW0002   AREA01   0.75
00AKGW0003   AREA01   0.33
00AKGW0003   AREA02   0.33
00AKGW0003   AREA03   0.33
```

This file indicates that the whole of 00AKGW0001 is to be allocated to new zone AREA01, a quarter of 00AKGW0002 goes to AREA02, with its remainder going to 00AKGW0002, and that 00AKGW0003 is to be equally divided between AREA01, AREA02, and AREA03.

To enable SASPAC to make use of a gazetteer file, the location of the three elements - existing zone ID, new zone ID, and scale factor (if present) - within the file, must be defined. This is done in the INPUT GAZETTTEER FILE command, which takes the following form:

```
INPUT GAZETTEER FILE NAME = filename   /
EXISTING ZONE COLS a TO b/
SCALE FACTOR COLS c to d /
NEW ZONE COLS e to f
```

where a and b indicate the start and end column locations of the existing zone code, c and d the start and end column locations of the scale factor, and e and f the same for the new zone code.

Note that there is no fixed order for the elements on the gazetteer file, *but the order within the Command File is mandatory, and must be as shown.*

The scale factor need not be present in all cases - in fact it may not be present at all - but if it is to be used, then its location must be specified in the command. If it is used, and it is not present on a line, SASPAC assumes that all of the area is to be included.

The existing zone and new zone fields may overlap, which allows users to make use of the hierarchic code structure of the standard areas to aggregate to higher areas. However aggregation of Census Area Statistics and Standard Tables is not recommended as a means of achieving higher-level statistics due to the Small Cell Adjustments that are present in the data. It is preferable to obtain the higher level data direct from the Census Offices.

The Gazetteer File described above merely allocates a **ZONEID** to the new area and does not allocate a label or name to the new zone. There is a variant of the Gazetteer File which allows this. The presence of the words **WITH LABELS** following **FILE** in the gazetteer command instructs SASPAC to expect the file to consist of two parts separated by a blank line. The second part of the file would be the allocation to new zones and would look exactly the same as that already described. The first part would

consist of a line for each new zone, where the line would have a new zone identity followed by a new zone label or name. Thus the Gazetteer file described previously would appear as:

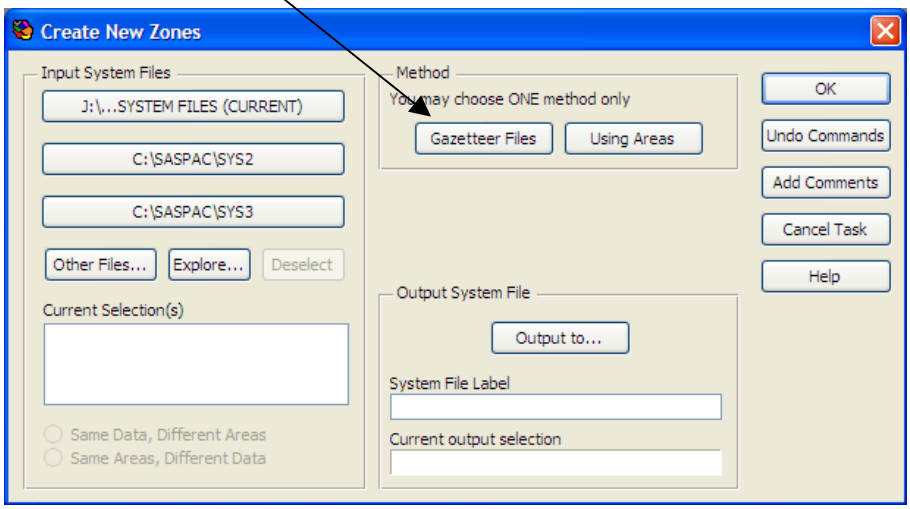
AREA01 This is area 1
AREA02 This is area 2
AREA03 This is area 3

00AKGW0001	**AREA01**	
00AKGW0002	**AREA02**	**0.25**
00AKGW0002	**AREA01**	**0.75**
00AKGW0003	**AREA01**	**0.33**
00AKGW0003	**AREA02**	**0.33**
00AKGW0003	**AREA03**	**0.33**

if it were to be read as a 'WITH LABELS' file.

Example: Create new areas using the gazetteer file ENFIELD.GAZ.

1. Again the task is to create new zones or areas, and so the initial part of the task creation is as in Module 7A. This leads us again to the following screen, where this time the '**Gazetteer Files**' Command Button is used.



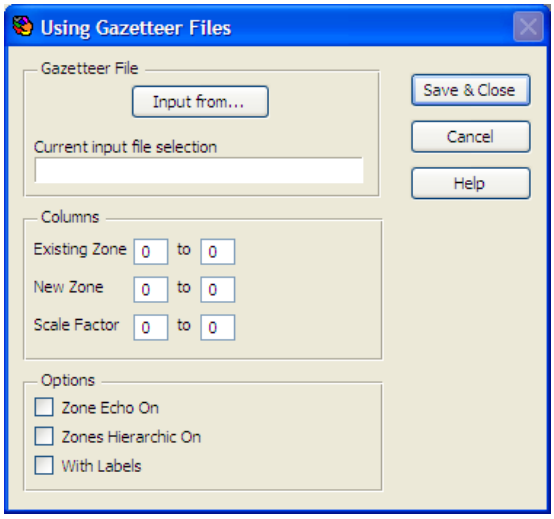
On use of this button the user is presented with a window where the various parameters of the gazetteer file have to be entered, and these (if not already known) have to be found from the gazetteer file itself. This may be done within SASPAC itself through use of the '**Open / Open Gazetteer File**' option, but only before the start of the task creation procedure. Once the procedure is commenced, the '**Open**' menu becomes unavailable. In this case, the file may be opened in Textpad, Notepad, or other such text editor, as available. If we open this gazetteer file in such a text editor, we see that its layout is as follows:

```
ward01  Bush Hill Park
ward02  Winchmore Hill
ward03  Enfield Town
ward04  Ponders End

00AKGT0001  ward01  0.33
00AKGT0001  ward02  0.67
00AKGT0002  ward01
00AKGT0003  ward03
00AKGT0004  ward02  0.5
00AKGT0004  ward03  0.5
00AKGT      ward04
```

2. The blank line following the assignment of labels to the new ward codes indicates that this is a gazetteer file with labels. It also shows that scaling factors are used and that they appear in columns 21 to 24. The existing zones are to be found in columns 1 to 10, while the new zones are in columns 13 to 18.

   Note also that the existing zones consist of both wards and OAs, and so both levels must be present on the input system files. With this information now available, we can proceed to the next stage of the procedure which produces the following window.

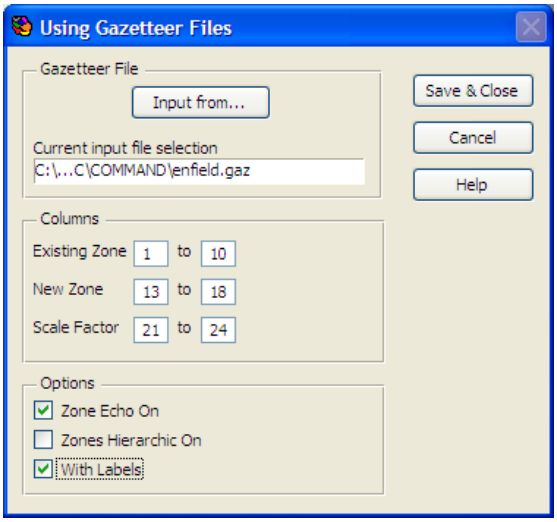


3. In this window the parameters discussed above may be entered. Additionally, there are a few check boxes available which should be considered.

Zone Echo On:	If this is checked, SASPAC will provide an expanded list of the allocation of existing areas (including scale factors) to new areas. This will appear in the Log file.
Zones Hierarchic On:	If the new areas are to nest into a hierarchic structure, then this box should be checked.
With Labels:	This box must be checked if the input gazetteer files has zone labels.

Following entry of the parameters in this window, including location of the gazetteer file, as shown below, the 'Close' Command Button is use to return to the main task window where the appropriate input system files are identified, as well as the location of the output system file.

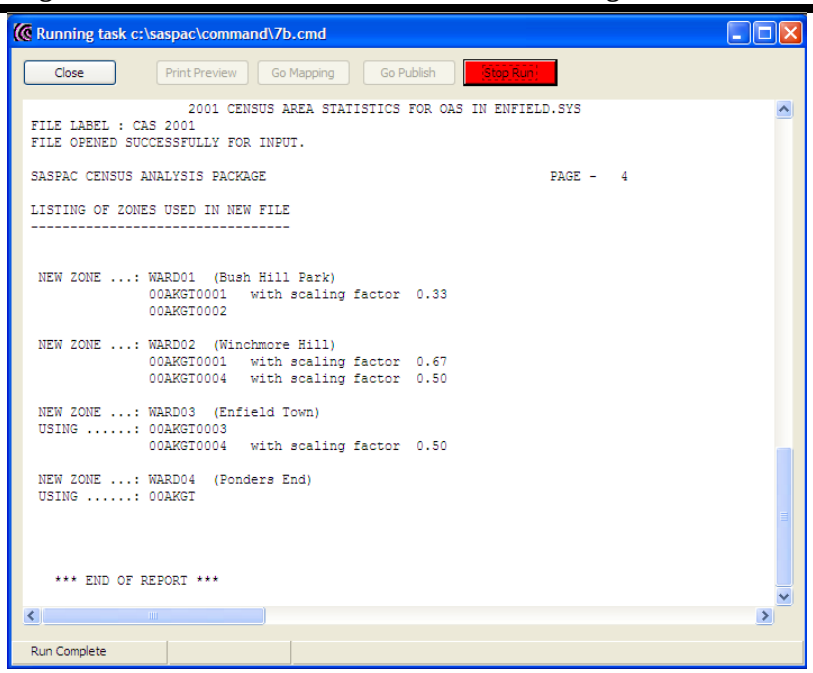
4. The standard procedure is then followed for running the task, including



allocating an user-defined name to the generated command file, the contents of which are as follows:

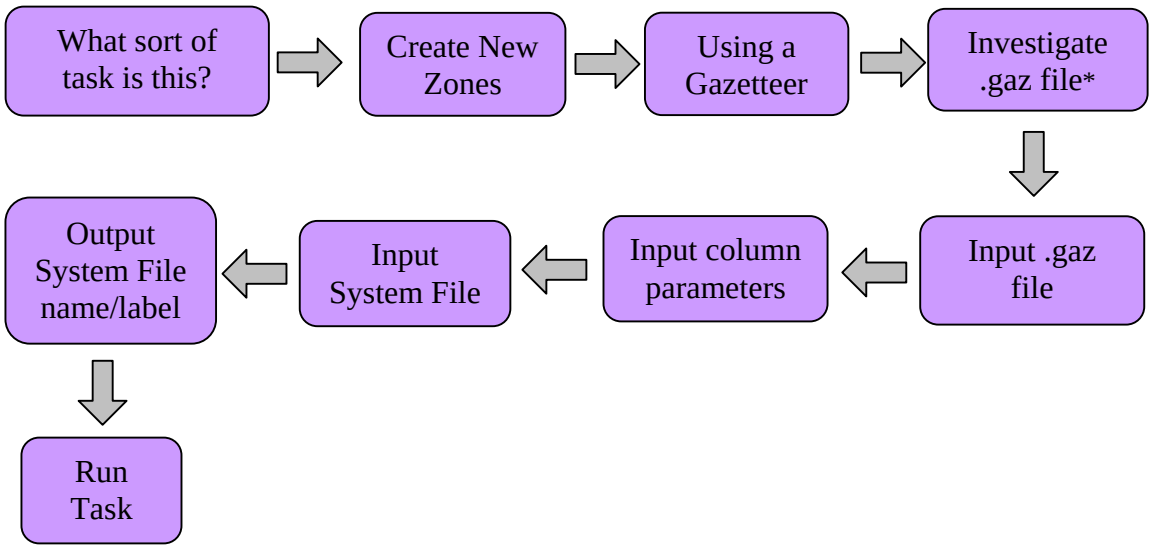
```
Input gazetteer file with labels name = /
C:\SASPAC\COMMAND\Enfield.GAZ existing zone cols
/
1 to 10 scale factor cols 21 to 24 new zone cols
/
13 to 18
input system file name = "C:\SASPAC\Training /
System Files\2001 CENSUS AREA STATISTICS FOR /
WARDS IN ENFIELD.SYS"
input system file name = "C:\SASPAC\Training /
System Files\2001 CENSUS AREA STATISTICS FOR OAS
/
IN ENFIELD.SYS"
read in series
set zone echo on
output system file name = /
C:\SASPAC\SYSFILES\TRAINING_MODULE_7B.SYS /
label=Training Module 7B Exercise
end
finish
```

The effect of using the 'Zone Echo On' parameter is shown in the following screen image of the Log File.


Running task c:\saspac\command\7b.cmd
Close
Print Preview
Go Mapping
Go Publish
Stop Run!
2001 CENSUS AREA STATISTICS FOR OAS IN ENFIELD.SYS
FILE LABEL : CAS 2001
FILE OPENED SUCCESSFULLY FOR INPUT.
SASPAC CENSUS ANALYSIS PACKAGE PAGE - 4
LISTING OF ZONES USED IN NEW FILE
NEW ZONE ...: WARD01 (Bush Hill Park)
00AKGT0001 with scaling factor 0.33
00AKGT0002
NEW ZONE ...: WARD02 (Winchmore Hill)
00AKGT0001 with scaling factor 0.67
00AKGT0004 with scaling factor 0.50
NEW ZONE ...: WARD03 (Enfield Town)
USING ....: 00AKGT0003
00AKGT0004 with scaling factor 0.50
NEW ZONE ...: WARD04 (Ponders End)
USING ....: 00AKGT
*** END OF REPORT ***
Run Complete

As with the previous example, there is no directly visible output from this task, and the only options are to run a task on the system file, or to investigate the system file' details.

5. A summary of the task sequence:

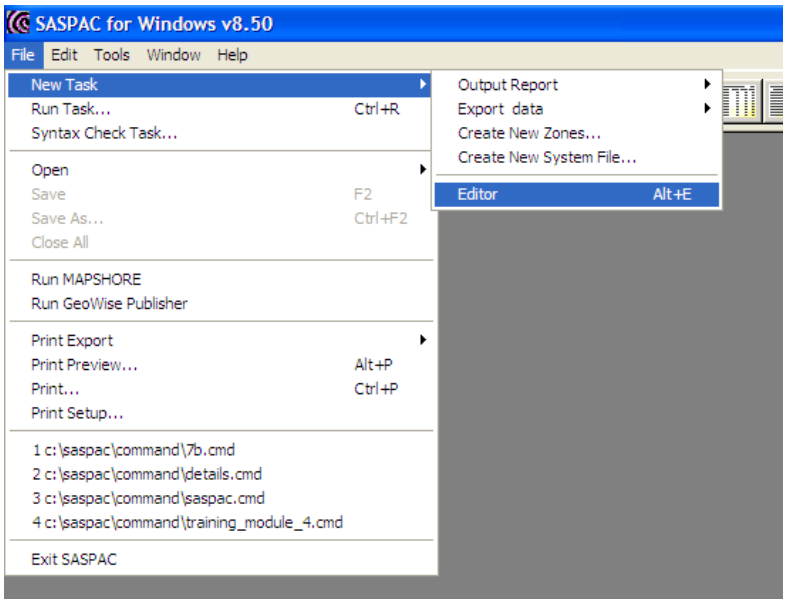


*.gaz file can be viewed within SASPAC(Open|Open Gazetteer file) or in a text editor package like notepad, textpad or Word.

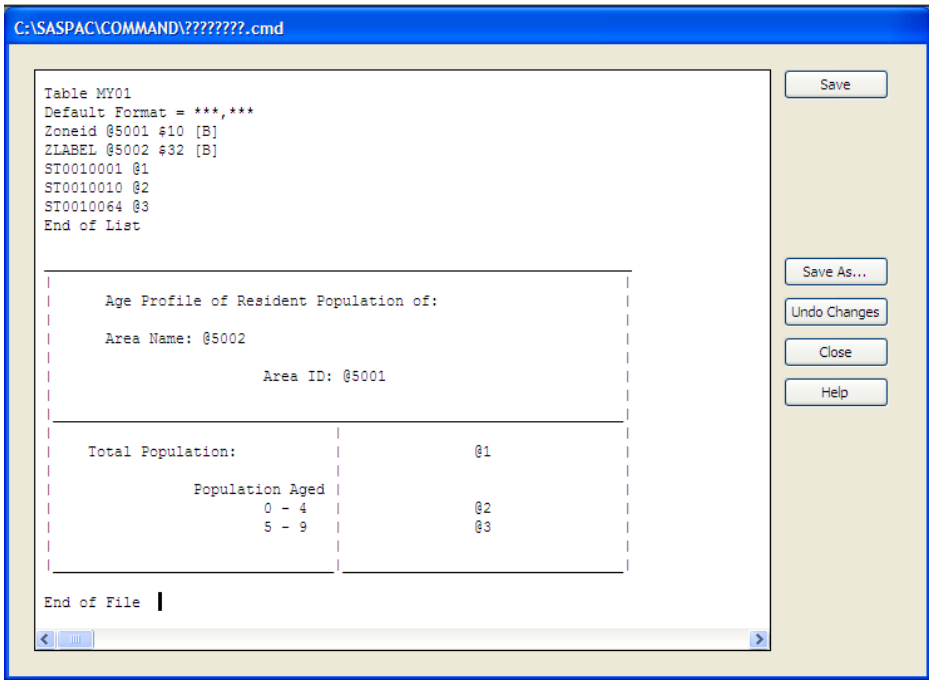
Module 8.

Creating User-Defined Tables.

1. SASPAC allows users to create their own table layouts, and in this module, we shall create a small table showing the age profile of an area. The task is initiated through the '**File / New Task / Editor**' option as shown.



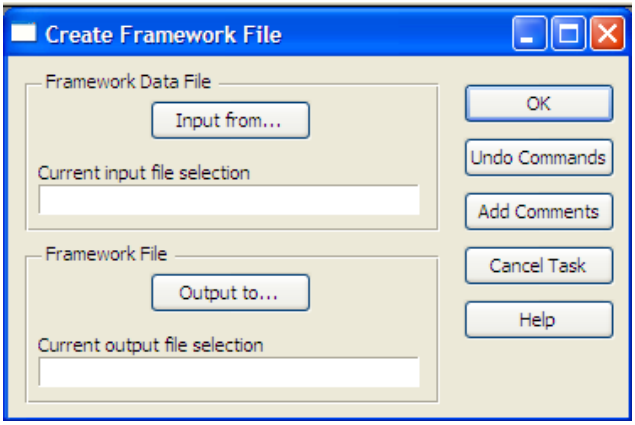
2. This opens up a blank window into which the user types the necessary text as below.



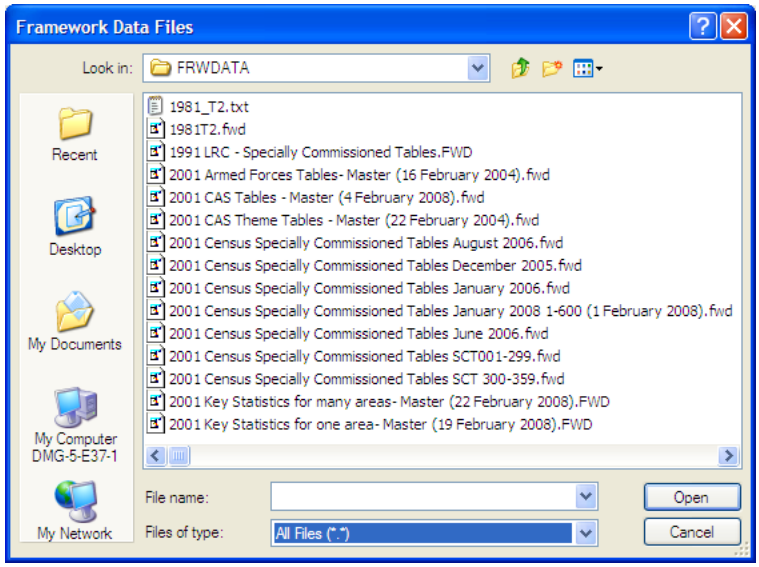
3. This file is saved to the FRWDATA directory through the standard Windows dialog box, as a **.CMD** file, as that is the only option for saving in an Edit box. This does not matter, because even though the framework file does not have the correct extension, it can still be used as one.

4. The next step involves converting the .FWD file into a .FWK file, through use of the '**Tools / Import Data / Framework File**' option

5. When this menu path is followed, the following window is presented to the user, and the source Framework Data File must be identified as well as the destination Framework File.



When the '**Input From**' Command Button is used, the system automatically defaults to the Framework Data directory as defined in the configuration file, and the file type automatically defaults to **.FWD**. The latter must be changed to '**All Files (*.*)**' before the required file - MYTABLE.CMD - can be selected. This is shown in the next screen image.


Framework Data Files
Look in:
FRWDATA
Recent
Desktop
My Documents
My Computer
DMG-5-E37-1
My Network
1981_T2.txt
1981T2.fwd
1991 LRC - Specially Commissioned Tables.FWD
2001 Armed Forces Tables- Master (16 February 2004).fwd
2001 CAS Tables - Master (4 February 2008).fwd
2001 CAS Theme Tables - Master (22 February 2004).fwd
2001 Census Specially Commissioned Tables August 2006.fwd
2001 Census Specially Commissioned Tables December 2005.fwd
2001 Census Specially Commissioned Tables January 2006.fwd
2001 Census Specially Commissioned Tables January 2008 1-600 (1 February 2008).fwd
2001 Census Specially Commissioned Tables June 2006.fwd
2001 Census Specially Commissioned Tables SCT001-299.fwd
2001 Census Specially Commissioned Tables SCT 300-359.fwd
2001 Key Statistics for many areas- Master (22 February 2008).FWD
2001 Key Statistics for one area- Master (19 February 2008).FWD
File name:
Files of type:
All Files (*.*)
Open
Cancel

Once the appropriate Framework Data file is selected, a name must be given to the output Framework file. Although there are no set rules on this, in practice, it is easiest to allocate the same root name to it.

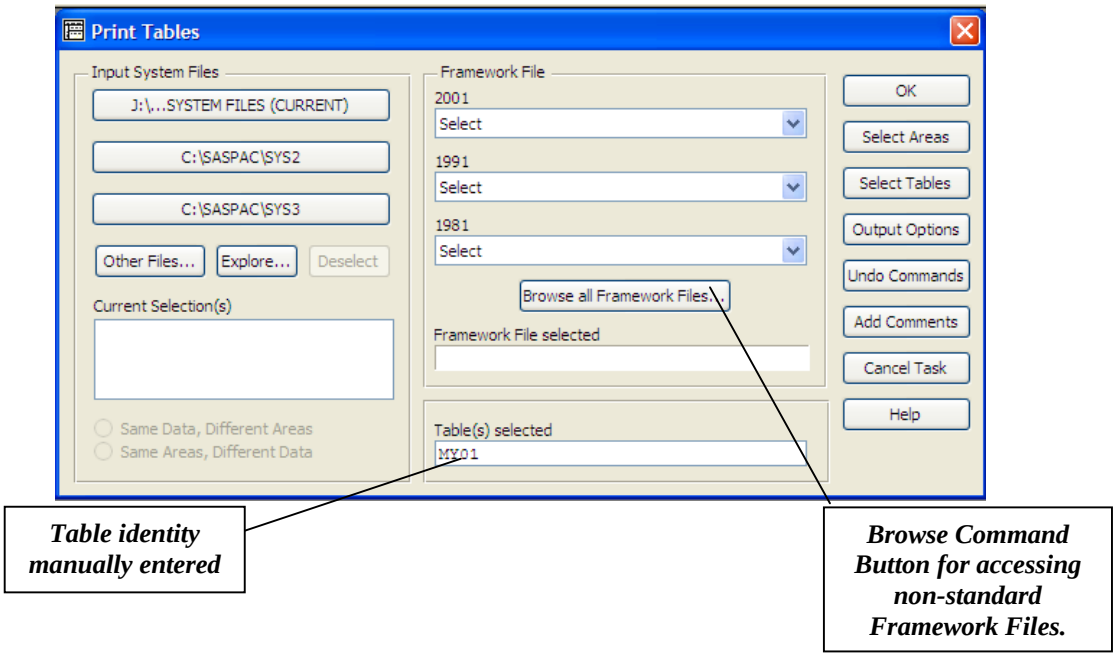
Occasionally there will be warning messages in this Log File, but as long as the message:

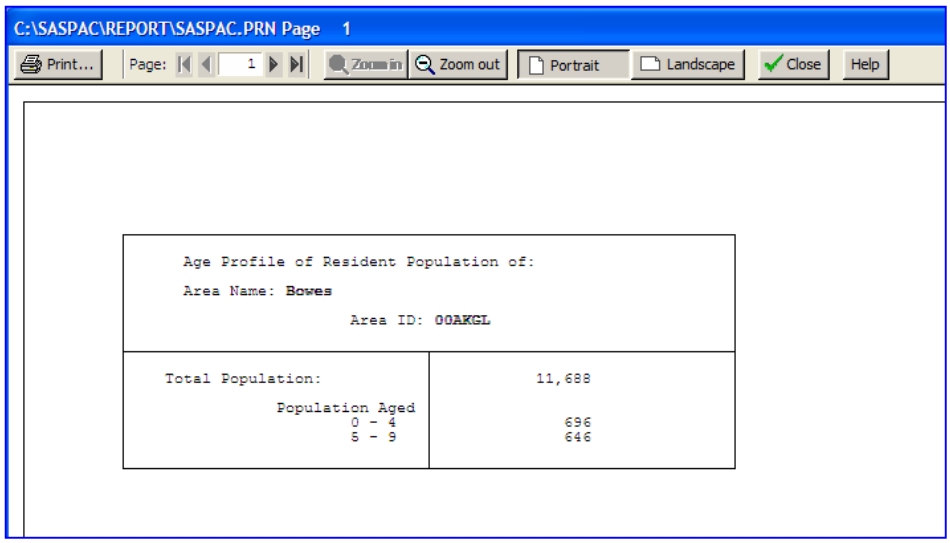
FRAMEWORK FILE CREATED SUCCESSFULLY

is obtained, there is no cause for concern.

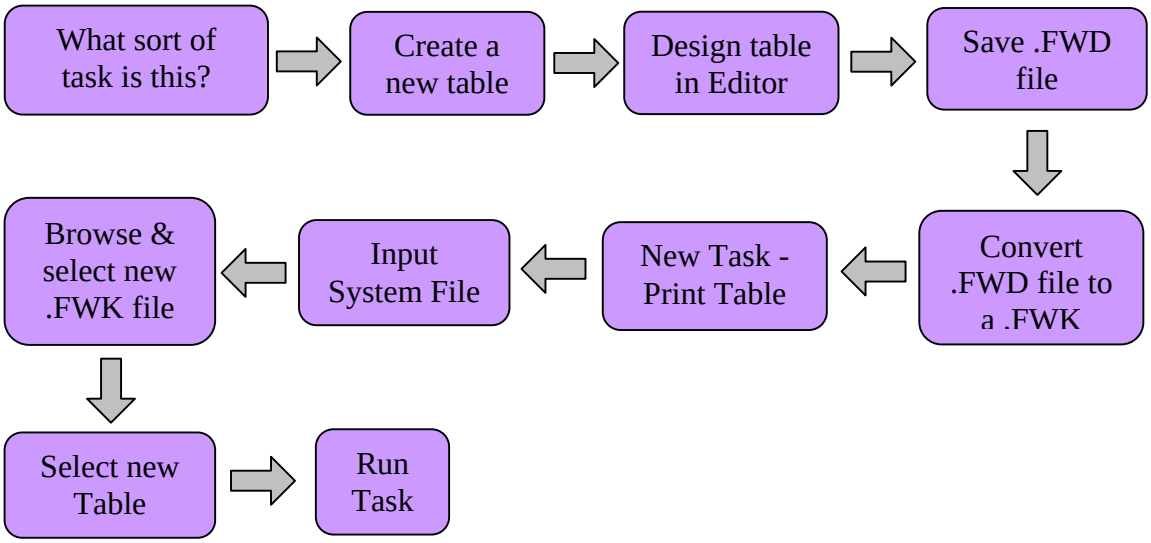
6. A standard 'Print Tables' task may now be created to view the results. In producing the Print Tables Command File, the Framework File is identified through use of the '**Browse All Framework Files**' Command Button, as it is a non-standard

file. Also the table identity cannot be selected as before, but it has to be manually entered in the appropriate text box. This is as shown on the following page, and the output generated is also shown there.





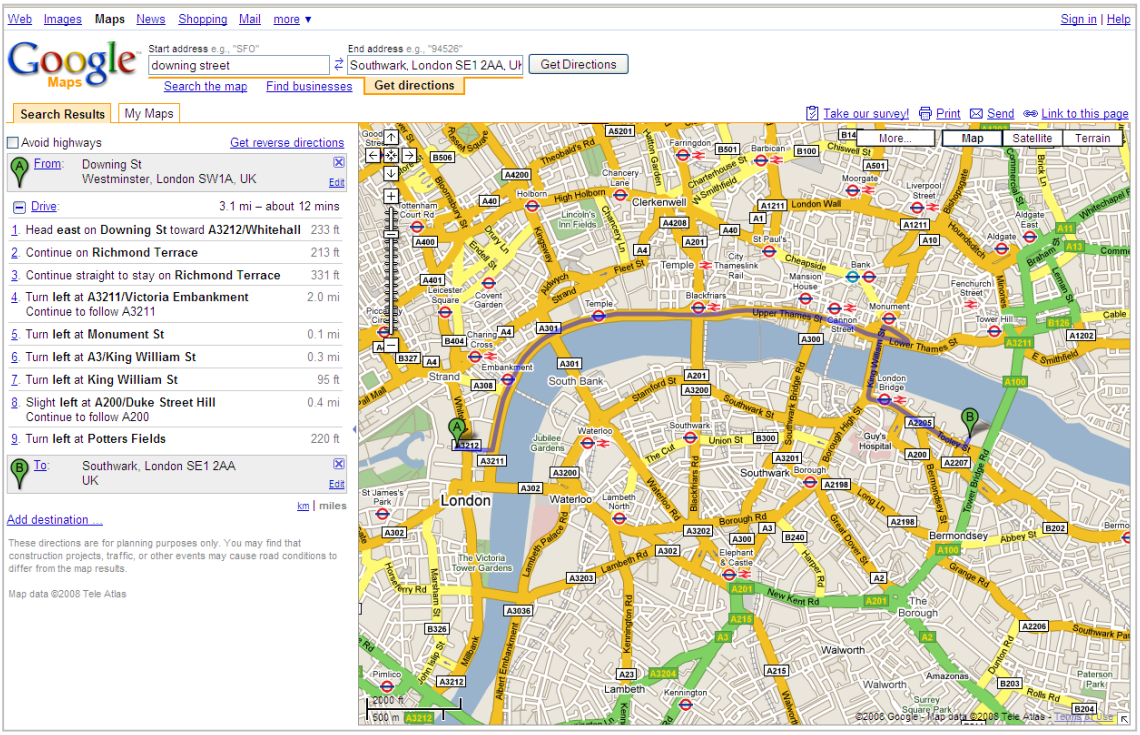
7. Summary of task sequence:


What sort of task is this?
Create a new table
Design table in Editor
Save .FWD file
Convert .FWD file to a .FWK
New Task - Print Table
Input System File
Browse & select new .FWK file
Select new Table
Run Task

Module 9.

Use of Google™ Maps within SASPAC.

The advent of Web Mapping has made mapping technologies and geographic data freely available to users on the World Wide Web. Previously, these facilities would be restricted to companies and mapping agencies with access to expensive hard and software, large quantities of geographic information and skilled cartographers. However, the release of Web Mapping applications like Google™ and Yahoo™ Maps have made these technologies accessible to everyone and greatly increased the use of these datasets. Currently, mapping applications typically include maps of the road network, aerial photography and routing tools (see screen shot below).

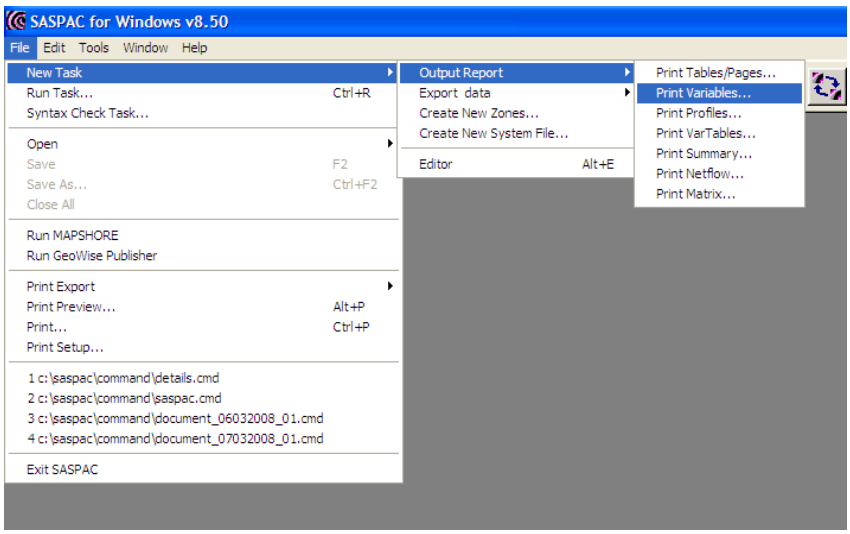


A screen shot of Google Maps data for Central London depicting a route between two locations as a blue line, with directions describing the route in the panel to the left of the map. It is possible to zoom in/out, scroll across the map, display aerial photography and terrain datasets.

In addition to providing access to these applications via a web address, many mapping providers make their services available for reuse in other web pages by publishing their API (Application Programming Interface). Since the release of SASPAC v8.5 users can take advantage of these technologies and seamlessly use Google™ Maps within a SASPAC task. The Google™ Maps facility can be launched through SASPAC's existing '**Select Areas'** and **'Radial Search**' function. This allows a user to search for a placename or postcode and visualise this on a map via www.google.co.uk. There are also a number of other SASPAC/Google™ Map utilities that are demonstrated through the following worked example.

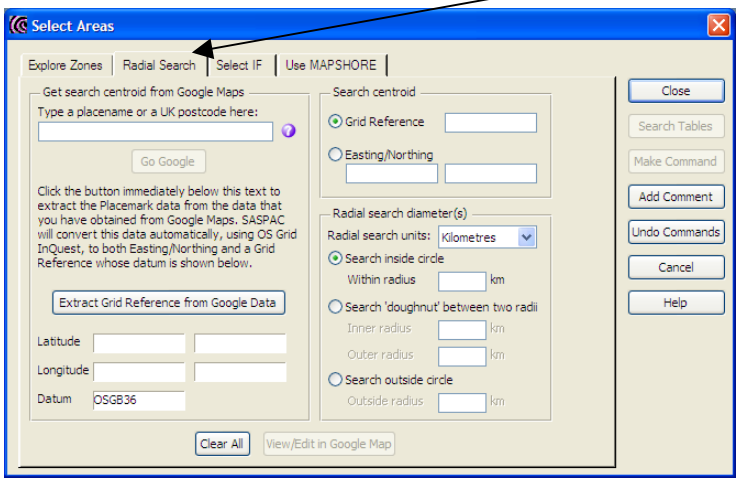
Example: Output, to a Print file, the population resident within 1 mile of London Bridge rail station, the population aged 16 to 74 in employment and the number of these that travel to work on foot. Use the 'Explore' tool to find the appropriate System file and Google™ Maps to locate and visualise the search area. Choose to print the file with statistics.

1. In this example, we are asked to define a specific search area, and output a number of variables to a print file. We must therefore define the task as one that outputs a report consisting of variables, and select areas using a radial search command. The usual procedure is thus followed, where the main task window is defined as '**File / New Task / Output Report / Print Variable**'.



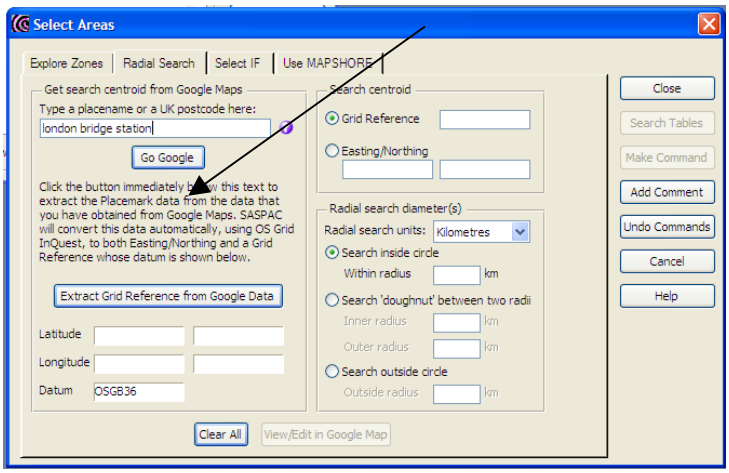
2. In the subsequent '**Print variables**' window there are a number of different options available to us. This example asks us to use the '**Explore**' tool to find the appropriate System file, it is advisable to define the variables required and select the areas that are within 1 mile of the search location first.

3. In this case we are asked to find the population resident within 1 mile of London Bridge station and to do so we need to select the Census Output Areas that fall within this zone. This is done through '**Select Areas**' and the '**Radial Search**' tab within the Print Variables task window.



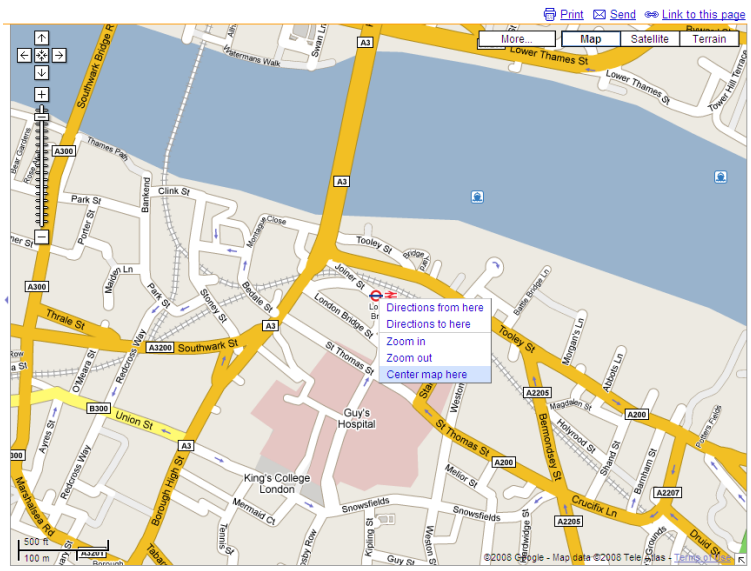
4. The '**Radial Search**' tab allows users to select areas selected based upon a defined distance from a fixed point or '**search centroid**'. This point can be defined by directly inputting, if known, an Ordnance Survey grid reference, or Easting and Northing, in the appropriate format. Alternatively, if this information is not know, a placename or UK postcode can be typed into the text box and Google™ Maps used to define a location.

5. This example asks us to select areas that are within 1 mile of London Bridge rail station and as we do not know the coordinates of this position, the first stage is to define the location of this point. To do this we type the text "london



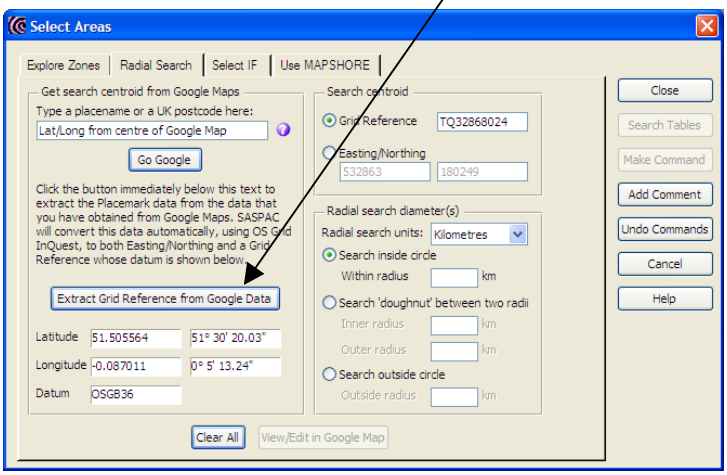
bridge station" into the text box and select '**Go Google**'

6. This opens an internet browser window, using Internet Explorer, with a red placemark over the approximate location of London Bridge station. In this case, the placemark is slightly southeast of the station itself, so to correct this we centre the map on the station by positioning the cursor over the station symbol and right-click the mouse button. Of the five options that are presented, choosing to '**centre map here**' will shift the centre of the map to the cursor position.

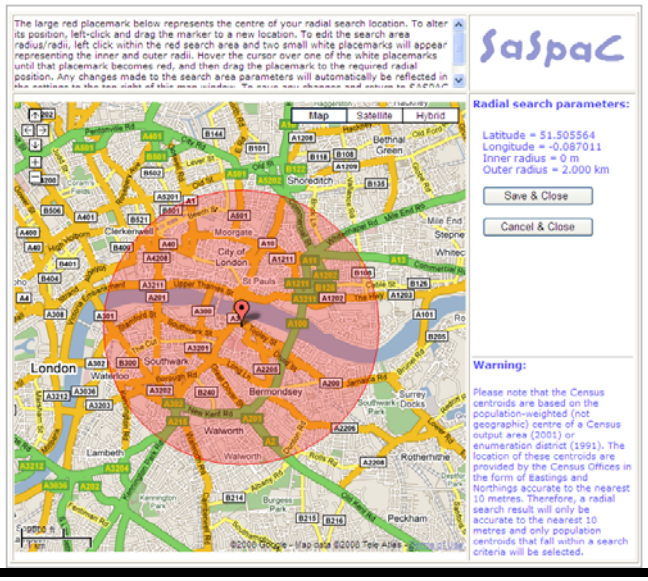


7. Now that we are happy with the map position we can send this location back to SASPAC using the '**Link to this page**' tool above the top right-hand corner of the map. This opens a '**Paste link in email or IM**' window and by right-clicking the mouse on the highlighted text (or 'CTRL-C' on your keyboard), the location will be copied to your PC's clipboard.

Now we can return to SASPAC and select the '**Extract Grid Reference from Google Data**' button to paste the geographical position into SASPAC. Notice that the latitude and longitude have been converted to the equivalent Ordnance Survey grid and easting/northing references.


Select Areas
Explore Zones
Radial Search
Select IF
Use MAPSHORE
Get search centroid from Google Maps
Type a placename or a UK postcode here:
Lat/Long from centre of Google Map
Go Google
Click the button immediately below this text to extract the Placemark data from the data that you have obtained from Google Maps. SASPAC will convert this data automatically, using OS Grid InQuest, to both Easting/Northing and a Grid Reference whose datum is shown below
Extract Grid Reference from Google Data
Latitude 51.505564 51° 30' 20.03"
Longitude -0.087011 0° 5' 13.24"
Datum OSGB36
Search centroid
Grid Reference TQ32868024
Easting/Northing 532863 180249
Radial search diameter(s)
Radial search units: Kilometres
Search inside circle
Within radius km
Search 'doughnut' between two radii
Inner radius km
Outer radius km
Search outside circle
Outside radius km
Clear All
View/Edit in Google Map
Close
Search Tables
Make Command
Add Comment
Undo Commands
Cancel
Help

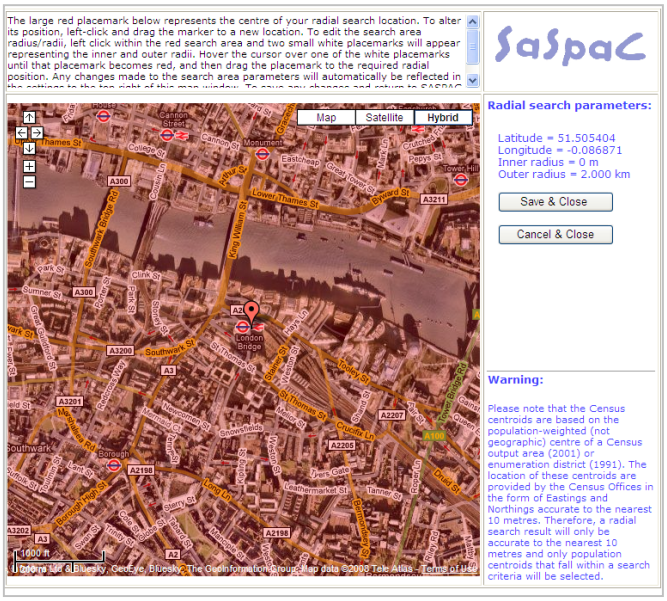
8. The next step is to define the radial search diameter, in this case we set the units to '**miles**' and type '**1**' into the '**Search inside circle**' text box. Having done this the '**View/Edit in Google Map**' button becomes available. It is recommended that you always check the location of the radial search on map before continuing. Selecting this button will open the following window:


The large red placemark below represents the centre of your radial search location. To alter its position, left-click and drag the marker to a new location. To edit the search area radius/radii, left click within the red search area and two small white placemarks will appear representing the inner and outer radii. Hover the cursor over one of the white placemarks until that placemark becomes red, and then drag the placemark to the required radial position. Any changes made to the search area parameters will automatically be reflected in
SaSpaC
Radial search parameters:
Latitude = 51.505564
Longitude = -0.087011
Inner radius = 0 m
Outer radius = 2.000 km
Save & Close
Cancel & Close
Warning:
Please note that the Census centroids are based on the population-weighted (not geographic) centre of a Census output area (2001) or enumeration district (1991). The location of these centroids are provided by the Census Offices in the form of Eastings and Northings accurate to the nearest 10 metres. Therefore, a radial search result will only be accurate to the nearest 10 metres and only population centroids that fall within a search criteria will be selected.
Map
Satellite
Hybrid
London
Southwark
Bermondsey
Walworth
Lambeth
Peckham
Rotherhithe
Shoreditch
Clerkenwell
City of London
Moorgate
Mile End
Stepney
Bethnal Green

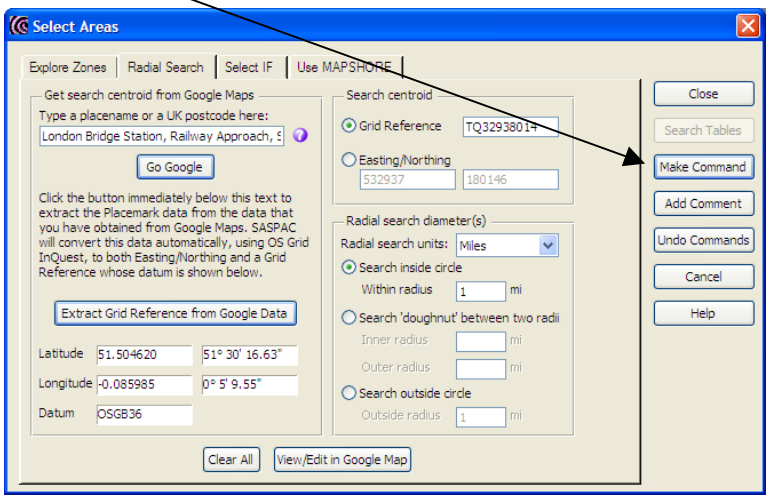
NB. Depending on the PCs internet security settings you may have to allow any blocked content to be displayed using the information toolbar above the web page.

This window correctly shows the radial search area as a circle with a 1 mile radius and a centre point positioned at London Bridge rail station.

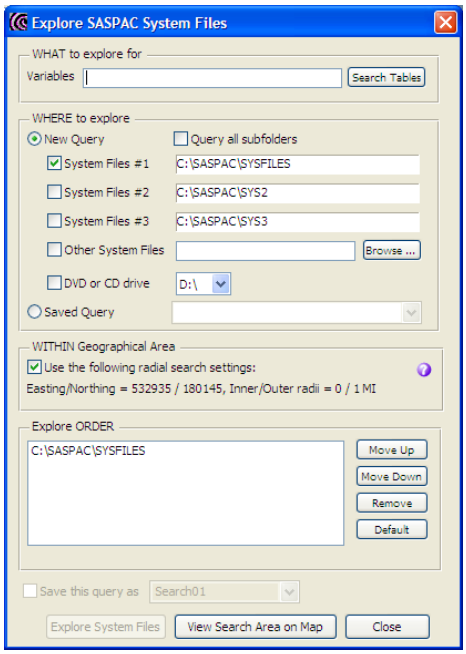
9. If necessary, it is possible to alter the position of the placemark and the radius of the circle following the instructions above the map window. We can also zoom in/out and choose to display aerial photography if desired, as shown below.


The large red placemark below represents the centre of your radial search location. To alter its position, left-click and drag the marker to a new location. To edit the search area radius/radii, left click within the red search area and two small white placemarks will appear representing the inner and outer radii. Hover the cursor over one of the white placemarks until that placemark becomes red, and then drag the placemark to the required radial position. Any changes made to the search area parameters will automatically be reflected in
SASPAC
Map
Satellite
Hybrid
Radial search parameters:
Latitude = 51.505404
Longitude = -0.086871
Inner radius = 0 m
Outer radius = 2.000 km
Save & Close
Cancel & Close
Warning:
Please note that the Census centroids are based on the population-weighted (not geographic) centre of a Census output area (2001) or enumeration district (1991). The location of these centroids are provided by the Census Offices in the form of Eastings and Northings accurate to the nearest 10 metres. Therefore, a radial search result will only be accurate to the nearest 10 metres and only population centroids that fall within a search criteria will be selected.

10. Now that we are happy with our search parameters we select the '**Save & Close**' button, to the right of the map, closing this window and returning to SASPAC. This returns us to the '**Select Areas**' window where we can choose to '**Make Command**' and save these parameters.


Select Areas
Explore Zones
Radial Search
Select IF
Use MAPSHORE
Get search centroid from Google Maps
Type a placename or a UK postcode here:
London Bridge Station, Railway Approach, S
Go Google
Click the button immediately below this text to extract the Placemark data from the data that you have obtained from Google Maps. SASPAC will convert this data automatically, using OS Grid InQuest, to both Easting/Northing and a Grid Reference whose datum is shown below.
Extract Grid Reference from Google Data
Latitude 51.504620 51° 30' 16.63"
Longitude -0.085985 0° 5' 9.55"
Datum OSGB36
Search centroid
Grid Reference TQ32938014
Easting/Northing
532937
180146
Radial search diameter(s)
Radial search units: Miles
Search inside circle
Within radius 1 mi
Search 'doughnut' between two radii
Inner radius mi
Outer radius mi
Search outside circle
Outside radius 1 mi
Clear All
View/Edit in Google Map
Close
Search Tables
Make Command
Add Comment
Undo Commands
Cancel
Help

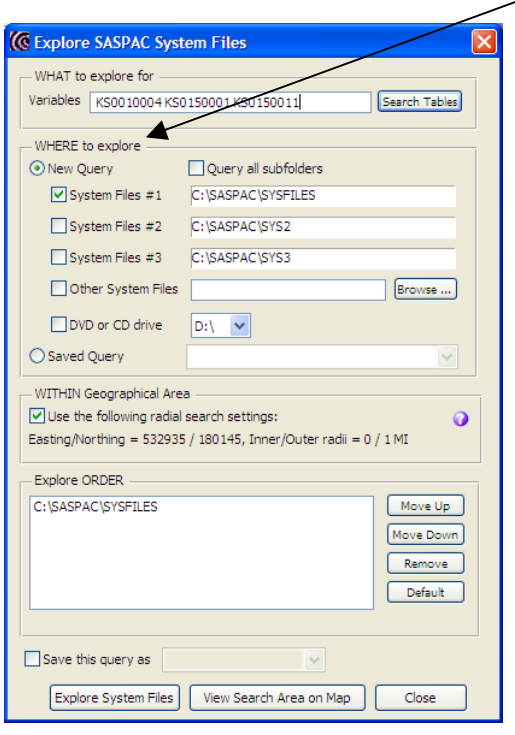
11. This returns us to the main task window where we can use the '**Explore**' button to search for and select the variables required. To do this we use the '**Explore**' and produce the following window:



12. The purpose of this window is to help us find the System File(s) that contain the data we are interested in. This may be defined as a series of variables with or without radial search parameters.

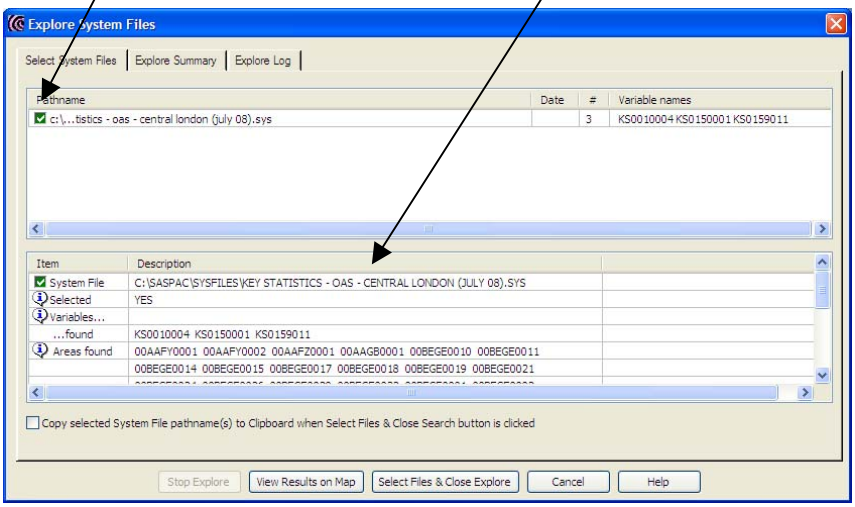
13. The variables of interest can be searched for in the usual manner under the '**WHAT to explore for**' heading. If the variable codes are known, these can be typed directly into the text box. Alternatively, the '**Search Tables**' button can be used to launch SASPAC's searching '**Select variables**' window and select the desired variables.

14. Now that we have defined the variables of interest, we specify where we would like SASPAC to search for this data under the '**WHERE to explore**' menu.



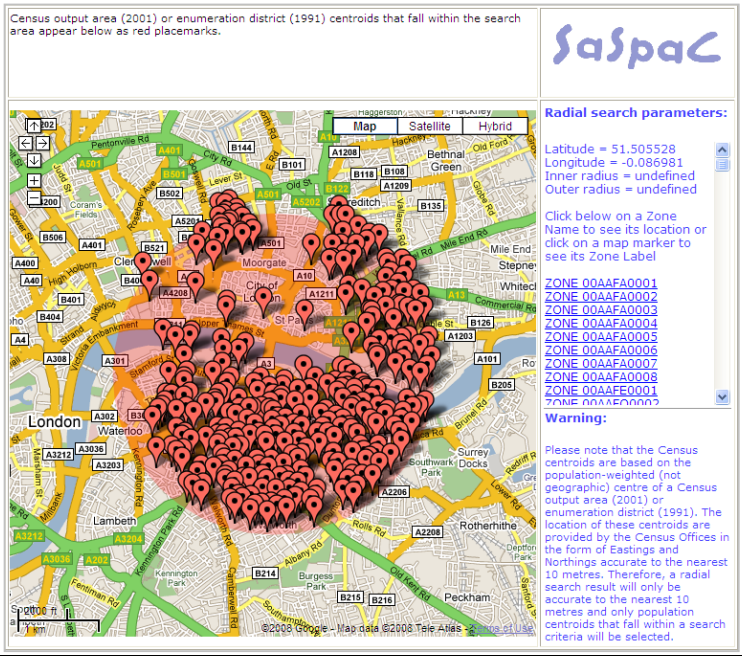
The default locations in this case are the standard System File directories (SYSFILES, SYS2 and SYS3) but it is also possible to '**Browse**' to a folder of choice or search a CD/DVD. In this case, the default search location of '**System Files #1**' is suitable. Before we run the search, note that because we have already defined radial search settings, the '**WITHIN Geographical Area**' box is checked and our parameters are displayed. To run the search we select the '**Explore System Files**' button at the bottom of this window.

15. After a short wait, the results will appear as a list of System files that contain some or all of the variables requested. To display the details of the files shown left-click the tick box next to the path name and file description is displayed in the window below.



NB. The only datasets to contain Ordnance Survey grid references from the 2001 Census are those at Output Area level. Therefore, when carrying out a radial search task, Output Area data must be used.

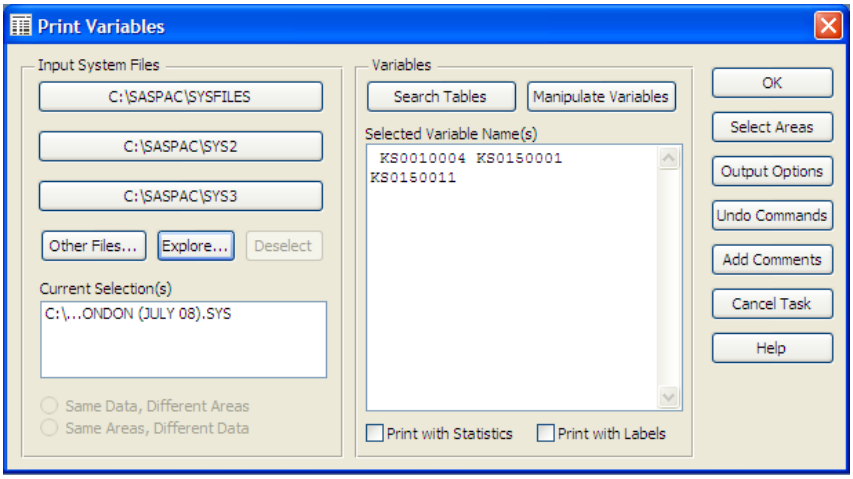
16. In this case only one System file should be found to contain all the variables requested and selecting the '**View results on Map**' button will visualise these radial search results on a Google™ Map:


Census output area (2001) or enumeration district (1991) centroids that fall within the search area appear below as red placemarks.
SASPAC
Map
Satellite
Hybrid
Radial search parameters:
Latitude = 51.505528
Longitude = -0.086981
Inner radius = undefined
Outer radius = undefined
Click below on a Zone Name to see its location or click on a map marker to see its Zone Label
ZONE 00AAFA0001
ZONE 00AAFA0002
ZONE 00AAFA0003
ZONE 00AAFA0004
ZONE 00AAFA0005
ZONE 00AAFA0006
ZONE 00AAFA0007
ZONE 00AAFA0008
ZONE 00AAFE0001
Warning:
Please note that the Census centroids are based on the population-weighted (not geographic) centre of a Census output area (2001) or enumeration district (1991). The location of these centroids are provided by the Census Offices in the form of Eastings and Northings accurate to the nearest 10 metres. Therefore, a radial search result will only be accurate to the nearest 10 metres and only population centroids that fall within a search criteria will be selected.
London
Waterloo
Lambeth
Rotherhithe
Peckham
Bethnal Green
Mile End
Whitechapel
Southwark Park
Surrey Docks
Burgess Park
Kennington Park
©2008 Google - Map data ©2008 Tele Atlas

The above map shows a red disc representing a 1 mile circle around London Bridge station. Red placemarks appear at the (population) centres of each Output Area that fall within the search radius. The OA codes are also listed to the right of the map and are dynamically linked to the placemarks, so when the zone ids are selected the associated placemark is identified.

NB. A radial search query includes an Output Area if the population's centre (centroid) is contained within the specified distance. For example, if only 25% of an Output Area is within a search radius but that happens to include the population centre, then that OA, and 100% of its population, is considered to be within that search area. If you would like to carry out more complex zone selections, i.e. zones 'completely' or 'partly within' search areas, we recommend using SASPAC's integrated mapping software - Mapshore.

17. Closing this window and returning to SASPAC, allows us to select the System File using the '**Select Files and Close Explore**' button. This identifies the selected file in the '**Current Selection(s)**' window and after choosing to '**Print with Statistics**', the task can be run using the '**OK**' button.



18. Navigating to the final page in the print file presents the following summary statistics for the data requested:

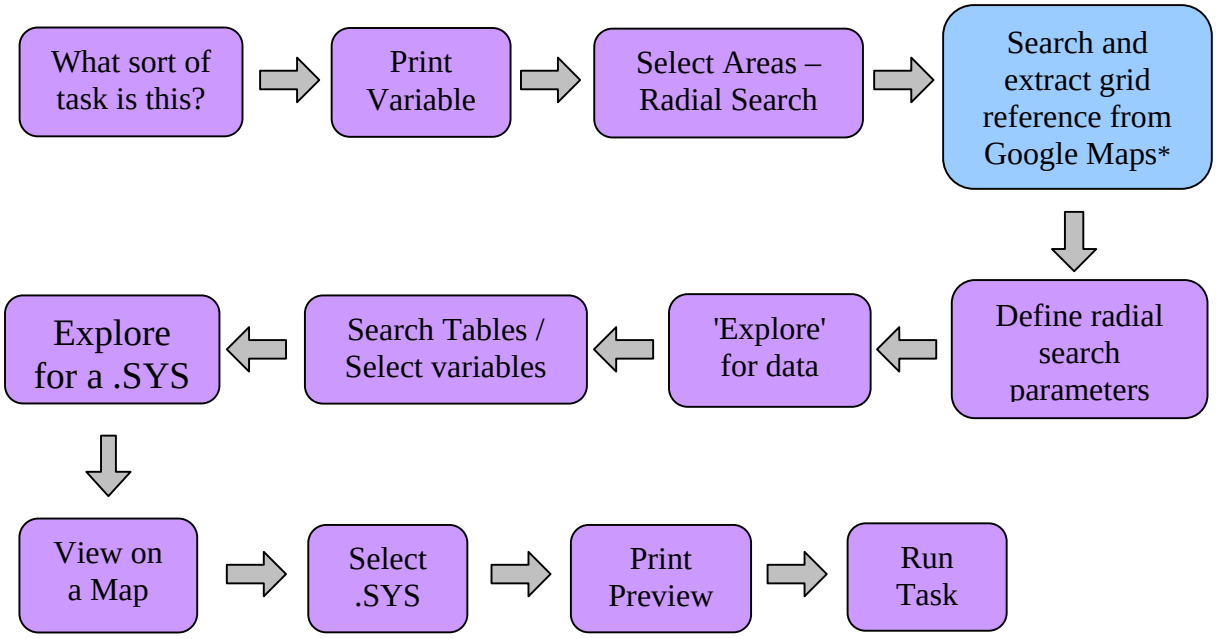
00BGGJ0023	172	115	38
00BGGJ0026	227	108	28
00BGGJ0029	282	88	22
00BGGJ0030	258	187	71
00BGGJ0031	268	147	52
00BGGJ0032	206	119	35
00BGGJ0034	233	181	71
00BGGJ0038	190	144	45
00BGGJ0041	225	170	67
00BGGL0003	301	108	29
00BGGL0018	197	123	45
00BGGL0019	303	133	40
00BGGN0012	342	152	39
00BGGN0015	199	167	50
00BGGN0019	232	182	60
TOTAL	57707	26690	6479
MEAN	276.11	127.70	31.00
STAND. DEV	113.40	44.24	19.49
MAXIMUM	1229	431	130
MINIMUM	106	57	6
MISS. VAL	0	0	0

This demonstrates that of the 57,707 people resident within 1 mile of London bridge rail station, 26,690 (or 46%) are aged between 16 and 74 and in employment and 9,939 (24%) of these travel to work on foot.

19. The command file generated by this sequence is as follows:

```
input system file name = "C:\SASPAC\SYSFILES\KEY /
STATISTICS - OAS - CENTRAL LONDON (JULY 08).SYS"
select if distmile(TQ32868024) le 1
* <<radial=1,532865,180245,I,mi,1,,>>
print variables with statistics  KS0010004 KS0150001 /
KS0150011
output print file = c:\saspac\report\Module9.prn
end
finish
```

20. Summary of task sequence:



* remember to 'Make Command'.

Module 10.

Use of Double Geography datasets within SASPAC.

Introduction

The 'standard' datasets supplied by the Census Offices are one-dimensional in a geographic sense in that each item of information is related to a single area. That is why they are occasionally referred to as **single geography** datasets. The CSV file as supplied consists of a series of records (or lines), each of which is a series of numbers associated with a single area code and (possibly) one area name. A part of the CSV file for table UV002 is shown here:

```
"00GANY0001",238,3,70.00
"00GANY0002",294,5,65.33
"00GANY0003",280,11,25.66
"00GANY0004",300,12,24.61
"00GANY0005",271,6,45.09
```

On the other hand, the 'special' datasets supplied by the Census Offices are geographically two-dimensional as each item of data is linked to two areas. In the migration datasets, the two areas are the area of residence one year prior to the census, and the area of residence at census day, while in the workplace (or in Scotland, travel) datasets, the areas are area of residence, and area of workplace (both at census day). These datasets are referred to as **double geography** datasets. As with the single geography datasets, the CSV file supplied consists of a series of records each of which is a series of numbers associated with two area codes and (possibly) two area names. A part of the CSV file for table SMS301 is shown here:

```
“20UEGL0006”,”20UEGF0017”,3,3,0,0,0,0,3,3,0,0,0,0
“20UEGL0006”,”20UEGF0019”,3,3,0,0,0,0,3,3,0,0,0,0
“20UEGL0006”,”20UEGG0006”,3,3,0,0,0,0,3,3,0,0,0,0
“20UEGL0006”,”20UEGG0008”,3,3,0,0,0,0,3,3,0,0,0,0
“20UEGL0006”,”20UEGJ0002”,7,4,3,0,0,0,7,4,3,0,0,0
“20UEGL0006”,”20UEGJ0003”,33,14,19,0,0,0,33,14,19,0,0,0
```

Apart from the fact that the double geography datasets have two geographic references per area, their record structure is similar to that of the single geography datasets, and SASPAC may be used to access them with a few minor modifications in the appearance of the windows and choices that are offered to the user.

For any single variable the dataset may be represented as a two-dimensional matrix – hence the alternative name matrix datasets – with the third dimension then being represented by the other variables.

The migration data is available as a United Kingdom dataset, but there are two datasets for workplace related movements. This is because in Scotland the question relating to place of work was extended to cover place of study as well. The two relevant questions in each of the areas was as follows:

England, Wales, and Northern Ireland:

Question 33. What is the address of the place where you work in your *main* job?

Question 34. How do you usually travel to work?

Scotland:

Question 10. What address do you travel to for your main job or course of study (including school)?

Question 11. How do you usually travel to your main place of work or study (including school)?

The Census Offices are releasing a series of tables under three generic titles:

Matrices of journeys to work (Areas of residence in England, Wales, and Northern Ireland)

Within SASPAC these tables are given the table identifier code SWS. There are 14 tables (plus one Northern Ireland variation). Seven of these – numbered SWS101 to SWS107 – are available at local authority district level; six – numbered SWS201 to SWS206 – are available at ward level; and one – numbered SWS301 – is available at Output Area level.

Matrices of journeys to place of work or place of study (Areas of residence in Scotland)

Within SASPAC these tables are given the table identifier code TVS. There are 14 tables. Seven of these – numbered TVS101 to TVS107 – are available at local authority district level; six – numbered TVS201 to TVS206 – are available at ward and postcode sector level; and one – numbered TVS301 – is available at Output Area level.

Matrices of migration moves

Within SASPAC these tables are given the table identifier code SMS. There are 16 tables (plus one Northern Ireland variation). Ten of these – numbered SMS101 to SMS110 – are available at local authority district level; five – numbered SMS201 to SMS205 – are available at ward level (and postcode sector level in Scotland); and one – numbered SMS301 – is available at Output Area level.

Within the United Kingdom there are some 220,000 Output Areas, and about 10,500 wards. The resulting matrices of area to area moves for each of these levels have the potential to be very large even though not all the combinations of origin zone and destination zone will actually produce a flow. The following table shows the number of origin zones, the number of destination zones, and the actual number of flows for a few of the datasets currently available.

Table	Number of origin zones	Number of destination zones	Number of flows
SWS201	**9,432**	**10,223**	**1,367,170**
SMS201	**10,610**	**10,608**	**851,317**
SWS301	**175,428**	**177,404**	**5,951,376**
SMS301	**221,468**	**221,817**	**1,971,282**

As table SWS301 for a single pair of origins/destinations contains 36 cells, the full dataset for this table contains over 200 million cells. A consequence of this is that the data is extremely sparse with most of the cells having a value of zero and many others having very low values. The Small Cell Adjustment Methods (SCAM)

applied by ONS and NISRA Census Offices to this data mean that a small count appearing in a table cell is adjusted - information on what constitutes a small cell count cannot be provided as this may compromise confidentiality protection – but basically any value of 1 or 2 is adjusted to become either a 0, or a 3, in such a way that loss of information is minimized. Totals and subtotals in tables are calculated as the sum of the adjusted data so that all tables are internally additive; within tables, totals and subtotals are the sum of the adjusted constituent counts. Because of this SCAMming, it has been found that the OA level matrices are unreliable, and must be used with great caution. ***The use of OAs as building blocks is therefore not advised for this dataset.***

As indicated, the datasets supplied by the Census Offices, and made available to SASPAC users as SASPAC System Files cover areas across the whole of the UK (or parts as applicable). This means that there will be flows on the files that users will have no interest in, and which will not be relevant to the user.

The full matrix can be divided into four component parts:

Flows within an area of interest
Flows into an area of interest from the rest of the country
Flows from an area of interest to the rest of the country
Flows totally external to the area of interest

and this is depicted schematically in the following diagram.

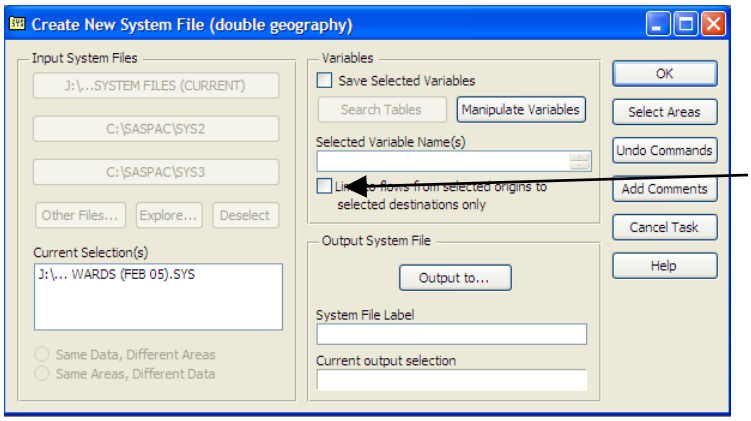
	Destination within area of interest	Destination outside area of interest
Origins within area of interest	Flows within area of interest **A**	Flows from area of interest to the rest of the country **B**
Origins outside area of interest	Flows into the area of interest from the rest of the country **C**	Flows wholly external to the area of interest **D**

Most users will only be interested in the shaded area within this diagram, and therefore the first requirement will be to reduce the dataset by removing the unshaded area. The area of interest in this case is termed the **Core Area**. The two blocks shown as **B** and **C** are referred to as the **Tails of the Dataset**.

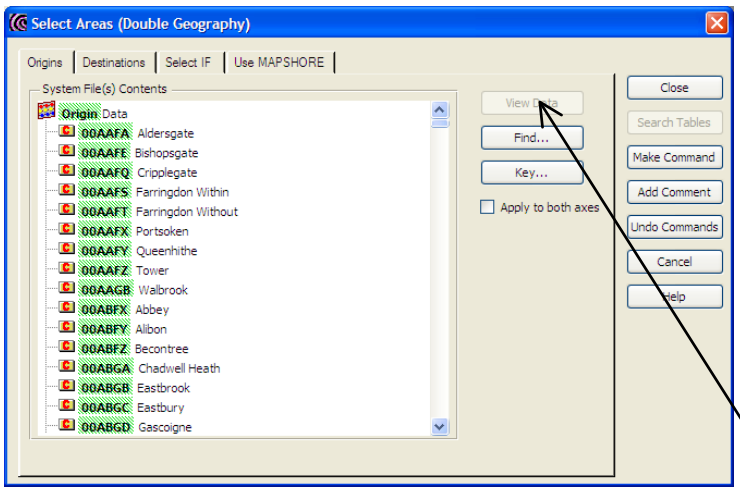
Creating a sub-set of the origin/destination datasets for zones within an area of interest (or Core Area).

For this exercise we shall take the Core Area as being the **London Borough of Enfield**, and we shall create the data for all Special Workplace Statistics Tables at ward level (2001 SWS2 WARDS TO WARDS.SYS).

The requirement is to create a new system file containing only data related to a core area, and so the New System File task window is accessed through use of the **File/New Task/Create New System File** option, at which point the appropriate national dataset is selected.



Once an origin/destination dataset is selected, the task window changes slightly in that there is an additional tick-box now available as shown in this image. This allows the user to restrict the core area data to only flows within the core area, and all flows with an external component will be excluded. In terms of the previous diagram, this would mean that only the moves in block A would be included. Above this tick-box under the 'Variables' header it is possible to input selected variables to be saved, leaving this blank will automatically save all variables in the original system file.

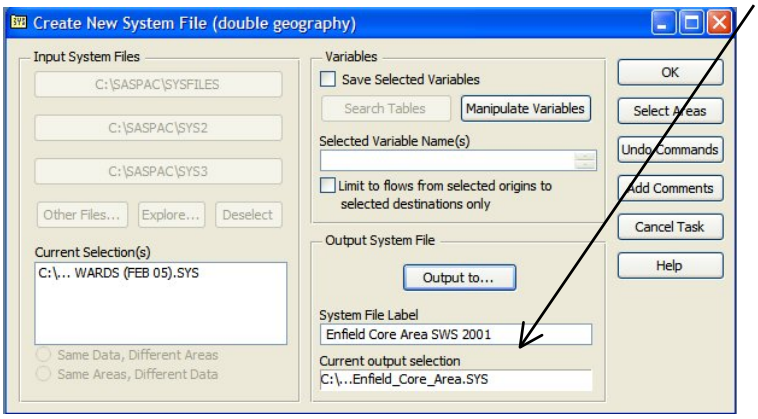


The selection of the core area zones may now be initiated in the usual way through use of the 'Select Areas' command button on the 'Create New System File' main task window. As a double geography file has been selected, the window that is now presented to the user reflects this in that the property sheet has tabs titled 'Origins' and 'Destinations'. To view the zones that are on the input file, the '**View Data**' command button is clicked, and a full list of the zones found on the input file will be shown. Note that all the area symbols will have a superimposed red 'C'. This indicates that the zone is part of the core area for this file. As it is a national data file all zones are within the core area and will therefore be indicated as such by the red 'C'

If the core area is to be symmetrical, that is the areas of origin within it are identical to the areas of destination, then the process of area selection may be simplified by clicking the 'Apply to both axes' tick-box. This ensures that the process of zone selection need only be done once. The first zone in the core area is selected by scrolling down the list and clicking on it followed by clicking on 'Include'. The list is then scrolled to the last zone in the area of interest, and by shift-clicking on this zone, the full range may be selected. ***Note however, that this may take some time, depending on the power of your PC and the number of zones to be included, and it may be faster to simply click on this last zone for inclusion and to edit the command file created as indicated below.***

[Note: An alternative to scrolling is to use the 'Find' button to access a search window into which the user enters a part of the code or name of the area that is to be used. In this case, as the code for ***Enfield*** *is* ***00AK****, we enter this in the text box, after which use of the 'Find next' button will move the selection to the first ward in* ***Enfield – 00AKGL Bowes****]*

Once the zones within the core area (or the start and end zones of the range) have been identified, then the 'Make Command' command button is clicked to generate the appropriate command(s). The final stage is to allocate an appropriate name and label in the required locations in the task window which appear when the 'Close' command button is clicked. Once this is done, the 'Close' command button on the task window is used to enter the task submission window, at which point the user can allocate an appropriate command file name. Also, if only the end zones of the range were selected, rather than the range itself, then the 'Save only' radio button on this window must be checked to allow the created command file to be edited as follows:



Change the lines which are of the form:

```
include <startzone_code>
include <endzone_code>
```

to

```
include <startzone_code> to <endzone_code>
```

This command file can then be saved and run to produce a system file containing only the flows relevant to the core area.

If the Tools/System File Details menu item is now used on this new file, the user will see that only the wards in their area of interest are shown as being part of the Core Area. Other areas will appear as origins and/or destinations, but without the core area indicator, they are automatically parts of the tails of the dataset.

The System File Details display also shows that there are now 20,194 flows within the dataset arising from 2,581 origins and 1,665 destinations. The file is therefore much more manageable as is also shown by the fact that the original file of 611 Mbytes has been reduced to less than 5.9 Mbyte.

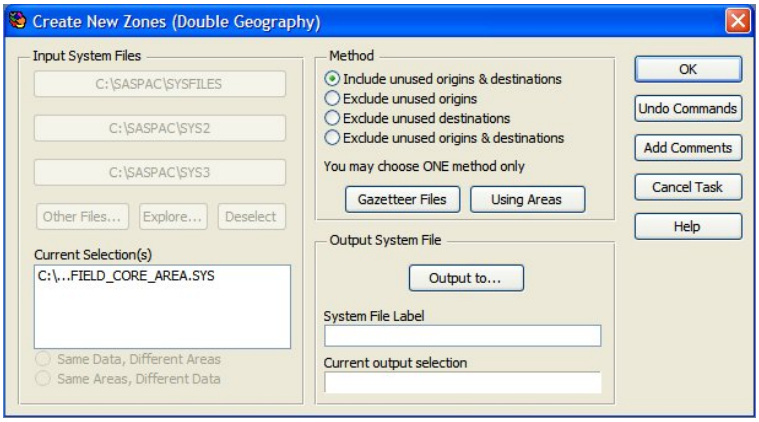
Re-zoning areas outside the core area.

The file that has just been created will contain ward-to-ward flows related to the **Enfield** core area. As such it will contain, for example, flows with origins in Scottish wards and destinations in **Enfield** wards. It is however highly probable that the user will not be interested in origins or destinations at such a fine level of detail where the distance involved is very large. In this case, the user will want to create new zones for areas outside the core area

The easiest form of external rezoning would see all the external wards aggregated to the local authority district level, although it would be possible to aggregate the more local wards to local authority level, those a little further away to county level, and those further still to region level.

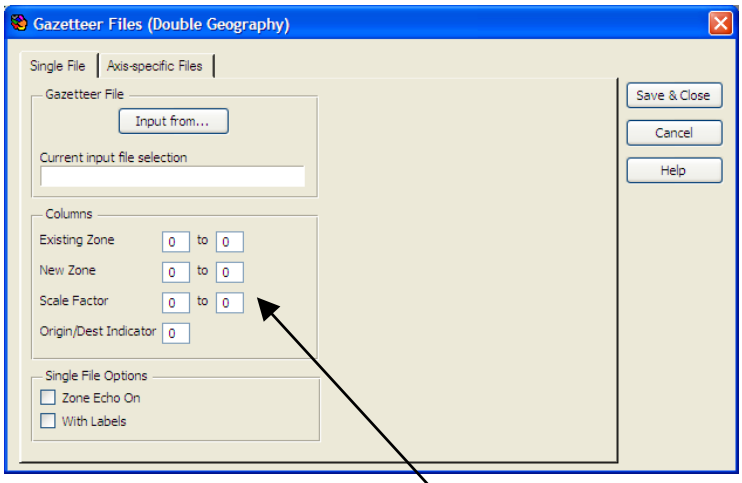
To undertake a rezoning of all external wards to local authority districts, we shall use the gazetteer file '**Ward_to_District.gaz**'. This file allocates wards to local authority districts and also provides labels for these districts. It has the existing zone (wards) in columns 1 to 6, and the new zone (districts) in columns 15 to 18, with no scaling factor. To adapt this for use in Enfield, we need to remove all wards in **Enfield** from the assignment (or second) part of the gazetteer file, and also the **Enfield** line from the labeling (or first) element of the file. Once this is done we will save the gazetteer file as '**Ward_to_District_Enfield.gaz'.**

The first stages of the process are the same as for single geography rezoning, and the "**File/New Task/Create New Zones..**" route is followed, at which stage the system file created in the previous exercise is selected. The resultant window now differs from the window seen in Module 7 in that this is double geography specific. There is a block under 'Method' which specifies what is to be done with areas on the input system file that are not used in the new zone creation commands – either through Using Areas or Gazetteer Files. As we are not considering the **Enfield** wards in the gazetteer file and we want those to appear as they are in the new system file, the 'Include unused origins and destinations' button is left checked.



As we are going to use a gazetteer file, the 'Gazetteer Files' button is used to take us to the next window, which is again different from that seen in Module 7, as it again

caters for double geography. The main difference is, that this window has two tabs which allow for gazetteer files to be chosen either to apply to both axes together or individually. This is shown in the next image. A minor difference is that as hierarchies are not allowed in double geographies, the 'Zones Hierarchic On' checkbox has disappeared.



The appropriate input gazetteer file is selected in the usual manner by use of the 'Input from…' button.

As noted earlier, the existing zone (wards) code is in columns 1 to 6, and the new zone (districts) code is in columns 15 to 18, with no scaling factor. These values are entered into the appropriate boxes. In addition, the gazetteer file has a label element, and so the 'With Labels' box must be checked. For the sake of obtaining a full check on the allocation of zones (wards) in the new zone creation, the 'Zone Echo On' box is also checked.

The final stage is to allocate a label and name to the output system file and then to run the rezoning task. **However, there is currently an issue that must be addressed before this is done and so the file must be saved and closed using the 'Save Only' option.**

As the gazetteer file has been generated from a listing of all ward files, and not all of these wards will appear as origins or destinations on the input system files. By definition, areas that appear on a gazetteer file are expected to appear on the input system file, and their absence will cause an error. This is overcome by use of a command that is used for testing the software, but in this case prevents the absence of the area from creating an error. This command is 'Test 71', and this is placed as the first command line in the command file.

A check on the System File details reveals that we now have a file in which there are 4,991 flows with 366 origin zones and 310 destination zones, and the file size has been further reduced from 5.9 Mbytes to 1.6 Mbytes.

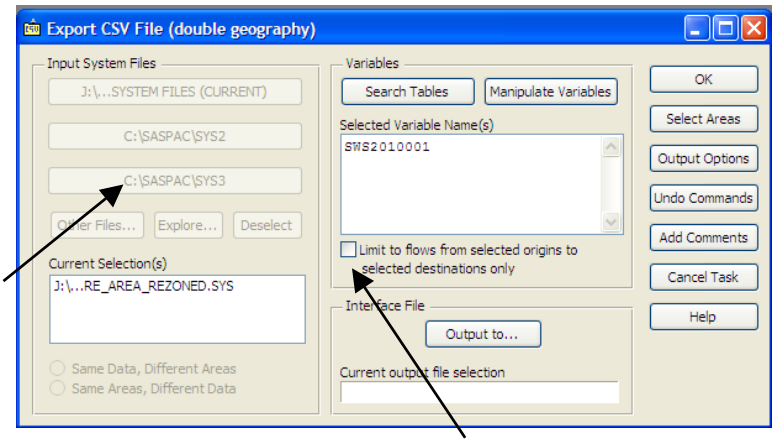
Using the origin/destination datasets.

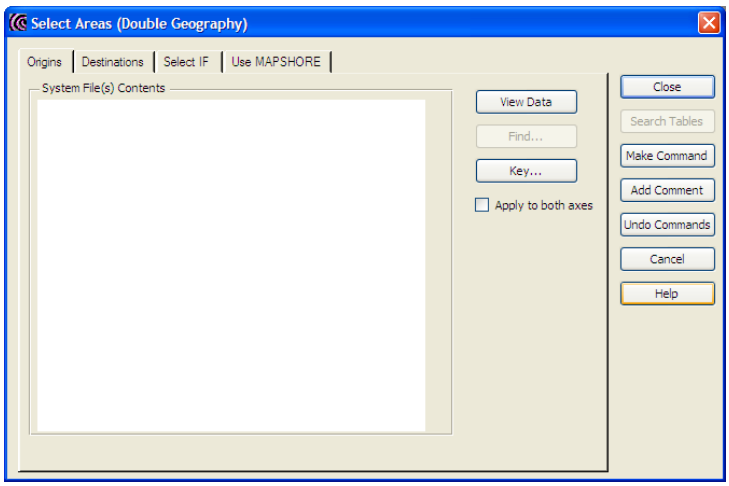
Once the national dataset has been tailored to the user's needs as detailed in the previous exercises, it may be used in virtually the same way as any other SASPAC System File. The main difference that the user will notice is that, once a double geography system file has been selected, some windows have a slightly different appearance, as has already been noted.

Example: For areas of origin and destination in London Borough of Enfield, extract the total numbers of persons, and males and females moving to work wholly within Enfield, and map the total persons within SASPAC.

This is a standard mapping task within SASPAC, with the difference that the dataset to be accessed is an origin/destination one. As there is a need to produce a CSV file for the mapping, the start of the task follows the usual **File/New Task/Export data/CSV file…** route. This produces the standard main task window where the user must select the required input file first. Note that whereas with single geography analyses, the input system file may be selected at any time, with double geography, the input system file must be selected at the outset. This is because the functionality available changes slightly when double geography is used.

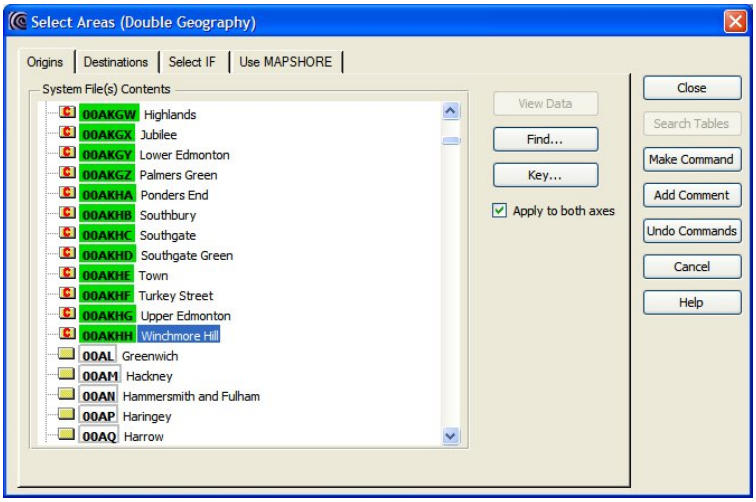
When the double geography file is selected, the amended window shown here will appear. Note that as soon as a double geography system file is selected, the 'Input System Files' buttons are greyed out. This is because only one input file is allowed with double geography analyses. Also note the addition of a check box that allows restriction of the analysis to flows wholly within the areas of interest. Checking this box leads to the insertion of the command 'SET IN&OUT OFF' in the Command File.





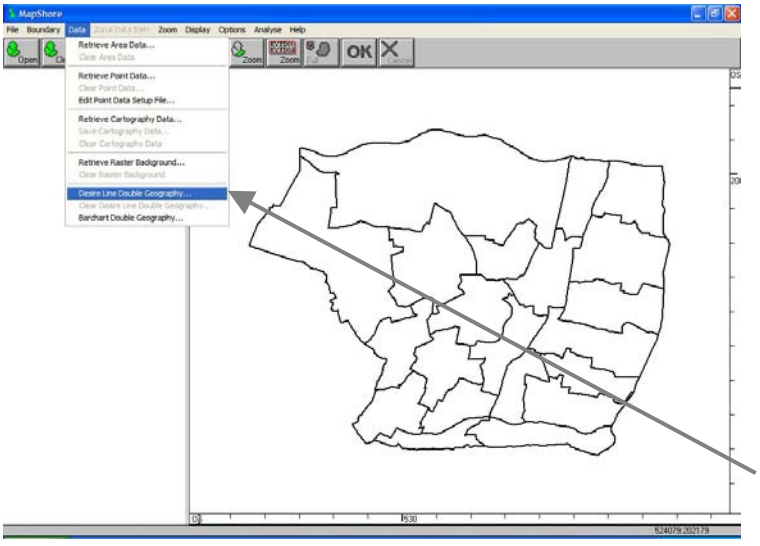
As with single geography, the 'Select Areas' command button may be used to select the required areas. However, as the input system file is double geography, radial searching is not an option, and in place of the 'Explore Zones' tab, there is a tab each for origin and destination.

Additionally, instead of the check boxes for the various exclusions associated with single geography, there is now a single check box labeled 'Apply to both axes'. If this box is checked, selections made in the usual way through either the 'View Data' or 'Use Mapshore' options will apply to both origin and destination. Otherwise, selections will have to be made twice - once for each of the origin and destination tabs.



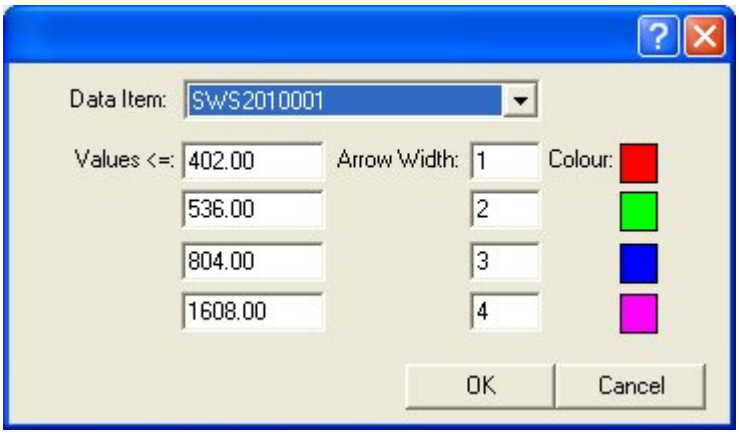
With the 'Apply to both axes' check-box checked, the areas within the core area may now be selected in the usual manner (*remembering to use the 'Make Command' button*). It is worth remembering that although the use of 'Shift/click' for selecting ranges works, sometimes in practice, depending on the size of the file, it can be a very slow procedure, and the most efficient way is to select the first area in the range, and then select the last area in the range. The commands created by this option may then be edited once the complete command file has been created, through use of the 'Save only' radio button in the final Task window.

The searching facilities for the double geography datasets can be used to search and select the required variables (NB. Remember to use the ward level SWS dataset to select these). Both the origin ('ORIGID') and destination'(DESTID') header variable codes will also need to be included at this point although, SASPAC will prompt the user to save these before running the task. Having done this, the output CSV file may be named and the task is run.

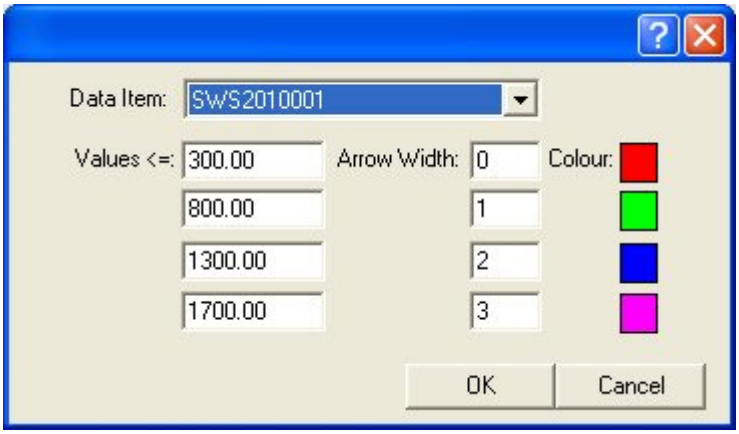


Using the 'Go Mapping' button in SASPAC will open Mapshore, and we then first open the appropriate boundary file – in this case **Enfield** wards – followed by the CSV file just created. As the CSV file is a double geography related file, the route to this is through the **Data/Desire Line Double Geography** option, as shown.

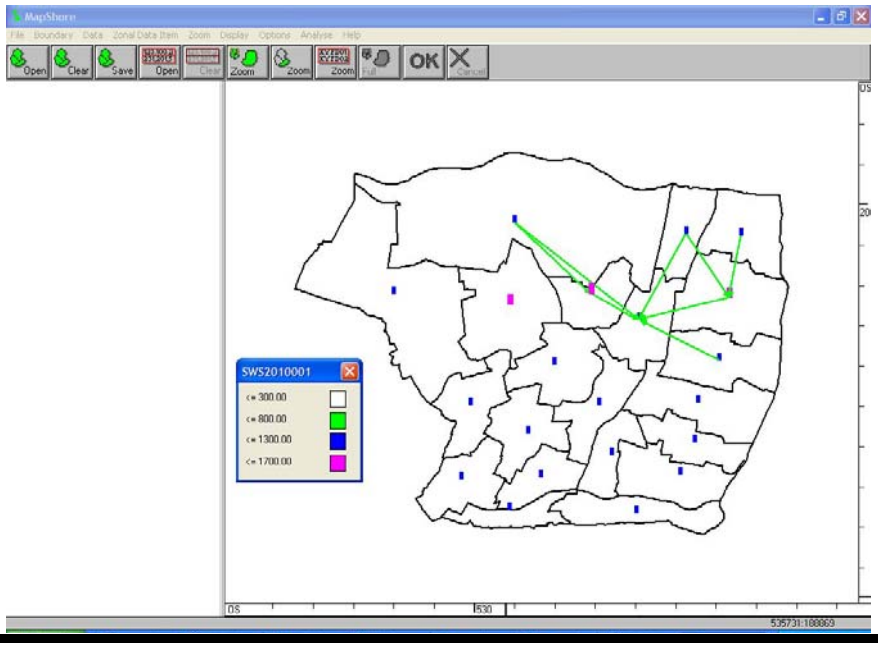
Use of this option will open the usual Windows dialog box where the user navigates to the appropriate drive/directory to find the required file. As there are multiple data items on this file that was created, the system asks which data item is to be mapped. If the variable representing the total (SWS2010001) is selected, the map is populated with a large number of arrows which effectively obliterate the whole map.
Use of **Zoom/Layering…** allows the user to control the presentation of the map.



This is done by highlighting the Double Geography file, and then using the 'Properties' command button. This brings up the window shown here, in which the user can set arrow widths, arrow colours, and range values. Setting an arrow width to zero will prevent that range from displaying on a map.



Using the settings as shown in the adjacent window, produces a map as shown below. This map can then be zoomed to examine flow patterns in particular areas. Note that because boundaries are not available for areas outside <AREA>, cross boundary flows would not be mapped even if they were present in the CSV file. The mapping depends on the availability of boundary files at the appropriate level.



As with any mapping within Mapshore, this can be annotated with text, and all the other mapping functionality is also available as discussed in Module 5.

Additional output facilities available only with Double Geography datasets.

In addition to the usual Print Tables and Print Variables options, there are three other Print options that are only available when a double geography dataset has been input. These are:

Print Summary
Print Netflow
Print Matrix

Additionally, it should be noted that Print Profile and Print VarTables are not available with double geography datasets.

Each of the three additional options is initiated in exactly the same way as the other Print commands.

Print Summary

A Print Summary task will print the following for each of the variables selected for output:

Row Total	The total number for the area as an area of origin, i.e. it is the sum of flows from that area of origin to all destinations.
Column Total	The total number for the area as an area of destination, i.e. it is the sum of flows to that area of destination from all origins
Intra	The total number of flows within the area, i.e. the area is the area of origin and destination
Row Max	The value of the largest flow from the area as an area of origin
Row Count	The total number of destinations with an origin in the area, and the number of these with a non-zero value for the selected variable
Column Max	The value of the largest flow to the area as an area of destination
Column Count	The total number of origins with a destination in the area, and the number of these with a non-zero value for the selected variable

Sample output from a Print Summary task is shown below.

SASPAC CENSUS ANALYSIS PACKAGE PAGE - 1

Variable: SWS2010001

Area Code	Row-Total	Col-Total	Intra	Row Max	Row Count	Col Max	Col Count
00AKGL	5720	1850	812	812	98/125	812	83/100
00AKGM	6539	2573	1207	1207	108/137	1207	68/101
00AKGN	5872	3517	1237	1237	102/128	1237	95/120
00AKGP	5474	4355	1110	1110	102/129	1110	125/166
00AKGQ	5170	6077	984	984	89/116	984	118/156
00AKGR	5961	8146	1439	1439	90/120	1439	127/188
00AKGS	5896	2295	994	994	94/128	994	80/102
00AKGT	5703	5056	1191	1191	99/128	1191	109/141

Print Netflow

A Print Netflow task will print the following for each of the variables selected:

Out flow	The value for the selected variable of the flow from the first area to the second area i.e. it is the value when the first area listed is considered as the origin, and the second area is considered as the destination
In flow	The value for the selected variable of the flow to the first area from the second area i.e. it is the value when the first area listed is considered as the destination, and the second area is considered as the origin
Ratio	The ratio of the Inflow to the Outflow
Net flow	The difference between the Inflow and the Outflow

Sample output from a Print Netflow task is shown below:

SASPAC CENSUS ANALYSIS PACKAGE PAGE -

Variable: SWS2010001

Origin	Destination	O to D Outflow	D to O Inflow	Ratio	NetFlo
00AKGL	00AKGL	812	812	1:1	0
00AKGL	00AKGM	12	23	0.52:1	-11
00AKGL	00AKGN	12	9	1:0.75	3
00AKGL	00AKGP	41	24	1:0.59	17
00AKGL	00AKGQ	55	18	1:0.33	37
00AKGL	00AKGR	42	9	1:0.21	33
00AKGL	00AKGS	9	9	1:1	0

Print Matrix

A Print Matrix task will print the matrix of moves between selected zones for the selected variables, with the areas of origin being the rows, and the areas of destination the columns.

Sample output from a Print Matrix task is shown below:

SASPAC CENSUS ANALYSIS PACKAGE PAGE -

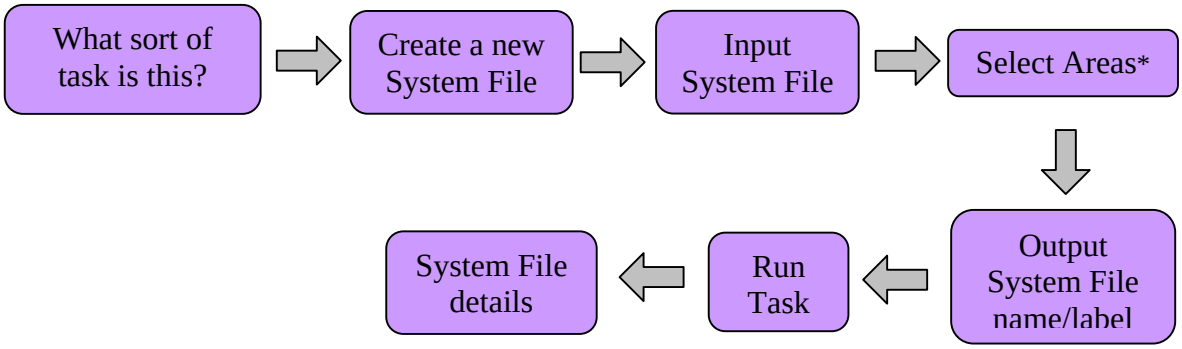
Variable: SWS2010001

Origin Dest	00AKGL	00AKGM	00AKGN	00AKGP	00AKGQ	00AKGR
00AKGL	812	12	12	41	55	42
00AKGM	23	1207	49	72	106	121
00AKGN	9	53	1237	49	79	164
00AKGP	24	18	11	1110	43	48
00AKGQ	18	36	35	27	984	126
00AKGR	9	43	89	35	152	1439
00AKGS	9	32	68	35	111	337

Note that the largest values are usually along the diagonal, as they represent the ‘within zone’ flows.

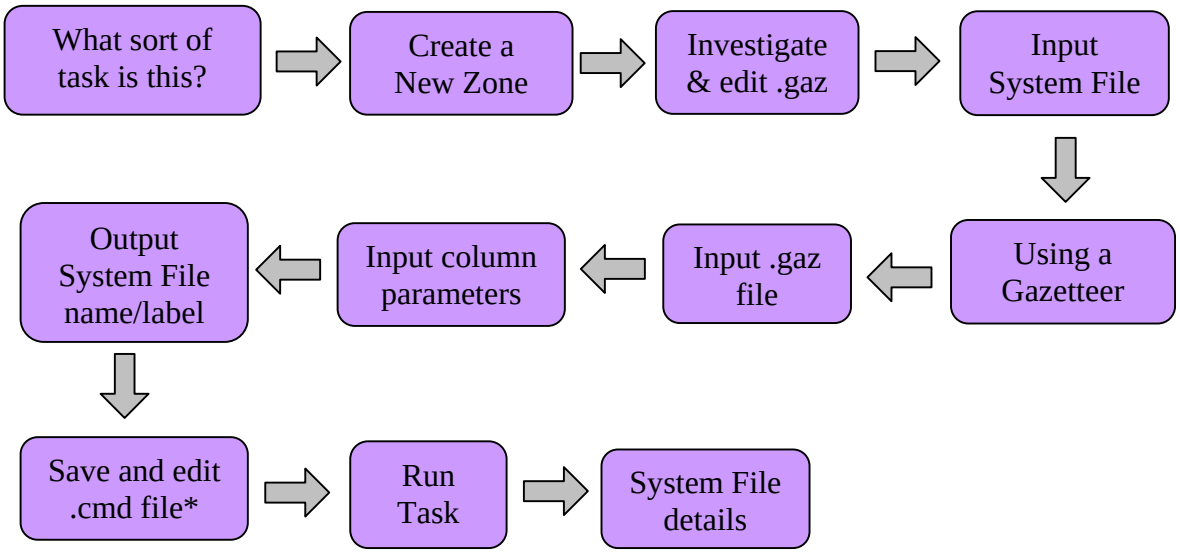
Summary of task sequence:

Step 1 – Create a subset of the origin/destination datasets:



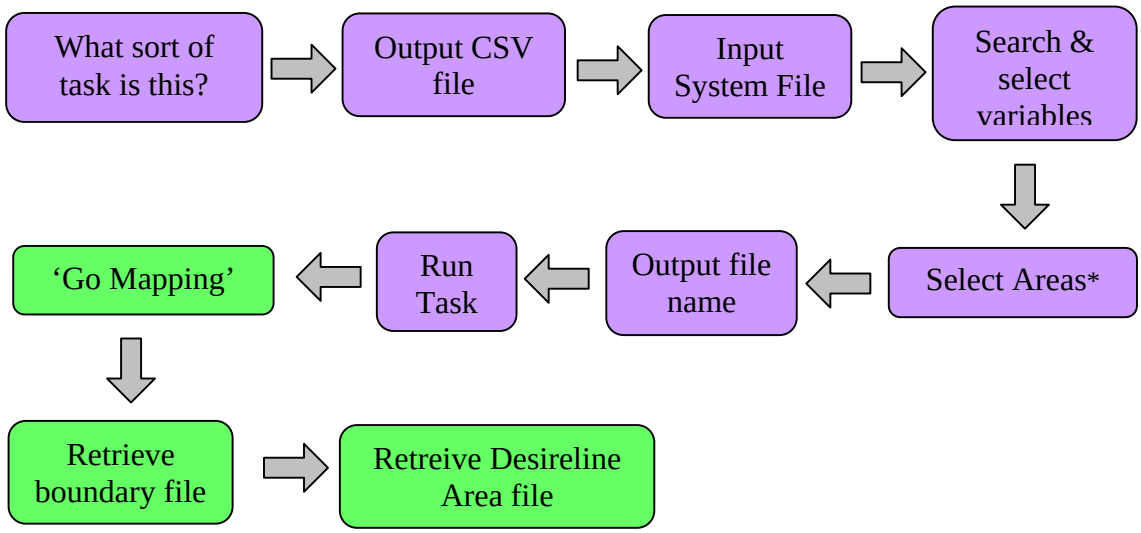
* remember to 'Apply to both Axes' as the selection is symmetrical and 'make command'

Step 2 – Rezoning areas outside the core area:



* insert double quotes around the .gaz file if there are any blanks in the directory or file name and the command 'Test 71'

Step 3 – Output from rezoned origin/destination file:



* remember to check the 'Apply to both axes' tick box to save making origin/destination selections seperately

Appendix 1

Worked examples answers

Appendix 1 - Trainers worked example answers:

Module 1

Command file:

```
input system file name = "C:\SASPAC\SYSFILES\/
2001 STANDARD TABLES FOR WARDS IN ENFIELD.SYS"
print variables ZLABEL ST0150049 ST0020255 /
ST0020270 ST0590031 ST0360229
output print file = C:\SASPAC\REPORT\ MODULE_1.PRN
end
finish
```

Print file:

SASPAC CENSUS ANALYSIS PACKAGE

ZONE ID	ZONE LABEL	ST0150049	ST0020255	ST0020270	ST0590031	ST0360229
00AKGL	Bowes	128	90	70	21	0
00AKGM	Bush Hill Park	141	101	125	109	6
00AKGN	Chase	127	99	115	81	0
00AKGP	Cockfosters	130	111	155	122	7
00AKGQ	Edmonton Green	94	99	109	15	0
00AKGR	Enfield Highway	112	89	90	66	6
00AKGS	Enfield Lock	103	58	76	50	3
00AKGT	Grange	170	122	139	99	3
00AKGU	Haselbury	88	94	87	38	6
00AKGW	Highlands	152	109	147	95	0
00AKGX	Jubilee	91	99	117	51	0
00AKGY	Lower Edmonton	69	95	80	29	0
00AKGZ	Palmers Green	149	102	101	51	3
00AKHA	Ponders End	108	80	70	33	5
00AKHB	Southbury	129	69	107	36	6
00AKHC	Southgate	178	88	143	55	6
00AKHD	Southgate Green	177	84	80	66	3
00AKHE	Town	144	104	150	72	0
00AKHF	Turkey Street	105	87	117	70	3
00AKHG	Upper Edmonton	52	89	91	23	3
00AKHH	Winchmore Hill	186	89	101	84	10

Module 2

Command file:

```
input system file name = "C:\SASPAC\SYSFILES\/
2001 STANDARD TABLES FOR DISTRICTS IN LONDON.SYS"
input framework file = c:\saspac\frwork\st01.fwk
print tables  ST052  ST103
output print file = C:\SASPAC\REPORT\MODULE_2.PRN
end
finish
```

Print file: *(Page 1 of 66)*

2001 Census Standard Tables — Area Identifier - **00AA** — Grid reference

PRODUCED USING SASPAC — **City of London** — CROWN COPYRIGHT RESERVED

Table ST052 TENURE AND PERSONS PER ROOM BY ACCOMMODATION TYPE
Table population: All Households

	ALL HOUSEHOLDS	House or Bungalow			Flat, maisonette or apartment			Caravan or other mobile or temporary structure
		Detached	Semi-detached	Terraced (including end-terrace)	In a purpose built block of flats	Part of a converted or shared house	In a commercial building	
ALL HOUSEHOLDS	4,346	16	11	72	3,868	132	247	0
Up to 0.5 persons per room	2,900	10	8	56	2,588	79	159	0
Over 0.5 and up to 1.0 persons per room	1,315	6	3	16	1,161	47	82	0
Over 1.0 and up to 1.5 persons per room	44	0	0	0	38	3	3	0
Over 1.5 persons per room	87	0	0	0	81	3	3	0
Owned	2,150	6	3	55	2,007	46	33	0
Up to 0.5 persons per room	1,579	3	3	46	1,470	37	20	0
Over 0.5 and up to 1.0 persons per room	535	3	0	9	501	9	13	0
Over 1.0 and up to 1.5 persons per room	12	0	0	0	12	0	0	0
Over 1.5 persons per room	24	0	0	0	24	0	0	0
Rented from Council	624	0	0	6	598	16	4	0
Up to 0.5 persons per room	386	0	0	3	373	6	4	0
Over 0.5 and up to 1.0 persons per room	217	0	0	3	207	7	0	0
Over 1.0 and up to 1.5 persons per room	6	0	0	0	3	3	0	0
Over 1.5 persons per room	15	0	0	0	15	0	0	0
Other social rented	318	0	0	0	300	14	4	0
Up to 0.5 persons per room	178	0	0	0	174	0	4	0
Over 0.5 and up to 1.0 persons per room	121	0	0	0	107	14	0	0
Over 1.0 and up to 1.5 persons per room	12	0	0	0	12	0	0	0
Over 1.5 persons per room	7	0	0	0	7	0	0	0
Private rented or living rent free	1,254	10	8	11	963	56	206	0
Up to 0.5 persons per room	757	7	5	7	571	36	131	0
Over 0.5 and up to 1.0 persons per room	442	3	3	4	346	17	69	0
Over 1.0 and up to 1.5 persons per room	14	0	0	0	11	0	3	0
Over 1.5 persons per room	41	0	0	0	35	3	3	0

Module 3a

Command file:

```
input system file name = "C:\SASPAC\SYSFILES\
2001 CENSUS AREA STATISTICS FOR OAS IN ENFIELD.SYS"
input framework file = c:\saspac\frwork\CS01.fwk
include 00AKGP0004
include 00AKGY0010 to 00AKGY0013
include 00AKHH0029
print tables CS001
output print file =  C:\SASPAC\REPORT\MODULE_3a.PRN
end
finish
```

Print file: *(Page 1 of 12)*

2001 Census CAS Tables (Census Area Statistics) Area Identifier - 00AKGP0004 Grid reference
Cockfosters Enfield
PRODUCED USING SASPAC ZONE 00AKGP0004 CROWN COPYRIGHT RESERVED

Table CS001 AGE BY SEX AND RESIDENT TYPE Page 1 of 2
Table population: All people

	ALL PEOPLE	Household Residents		Communal Establishment Residents	
		Males	Females	Males	Females
ALL PEOPLE	213	80	133	0	0
0 to 4	6	0	6	0	0
0	0	0	0	0	0
1	3	0	3	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	3	0	3	0	0
5 to 9	9	3	6	0	0
5	3	0	3	0	0
6	3	3	0	0	0
7	0	0	0	0	0
8	0	0	0	0	0
9	3	0	3	0	0
10 to 14	3	0	3	0	0
10	0	0	0	0	0
11	0	0	0	0	0
12	3	0	3	0	0
13	0	0	0	0	0
14	0	0	0	0	0
15 to 19	3	0	3	0	0
15	0	0	0	0	0
16	3	0	3	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0

Module 3b

Command file:

```
input system file name = "C:\SASPAC\SYSFILES\
2001 CENSUS AREA STATISTICS FOR OAS IN ENFIELD.SYS"
include 00AKHE0044
include 00AKGT0007
include 00AKHB0031
include 00AKHE0041
include 00AKHE0040
include 00AKHE0028
include 00AKHE0036
include 00AKHE0043
include 00AKHE0037
include 00AKHE0039
include 00AKHB0030
print variables with statistics CS0010001 CS0210010
output print file = C:\SASPAC\REPORT\MODULE_3b.PRN
end
finish
```

Print file:

```
 SASPAC CENSUS ANALYSIS PACKAGE

ZONE ID          CS0010001    CS0210010
00AKHE0044             296          175
00AKGT0007             260          167
00AKHB0031             207          108
00AKHE0041             309          148
00AKHE0040             214          100
00AKHE0028             374          166
00AKHE0036             333          169
00AKHE0043             309          148
00AKHE0037             317          155
00AKHE0039             300          150
00AKHB0030             254          101
```

Module 4

Command file:

input system file name = "C:\SASPAC\SYSFILES\/
2001 CENSUS AREA STATISTICS FOR OAS IN ENFIELD.SYS"
IF CS0010001 = 0 then percent1524 = 0 else percent1524=
(CS0010096+CS0010126) / CS0010001*100.0
describe variable percent1524 dp= 2 label=percentage aged 15-24
header all Percentage Population 15-24 OAs in Enfield
print variables with labels cs0010001 percent1524
output print file = c:\saspac\report\MODULE_4.prn
end
finish

Print file: *(Page 1 of 17)*

```
 SASPAC CENSUS ANALYSIS PACKAGE

Percentage Population 15-24 OAs in Enfield

ZONE ID      CAS Table CS   percentage
             001 CS001000   aged 15-24
00AKGL0001            273        11.36
00AKGL0002            199         6.03
00AKGL0003            304        13.16
00AKGL0004            311        15.11
00AKGL0005            234         5.98
00AKGL0006            284        13.73
00AKGL0007            248        11.29
00AKGL0008            255        10.59
00AKGL0009            375        15.20
00AKGL0010            307        15.96
00AKGL0011            226        11.50
00AKGL0012            357        11.20
00AKGL0013            273        10.62
00AKGL0014            328         7.62
00AKGL0015            299        11.04
00AKGL0016            260         8.46
00AKGL0017            313        15.02
00AKGL0018            337        14.54
00AKGL0019            346        14.74
00AKGL0020            392        17.09
00AKGL0021            313        12.14
00AKGL0022            277         8.66
00AKGL0023            310        14.19
00AKGL0024            289        19.38
00AKGL0025            355        11.27
```

Module 5

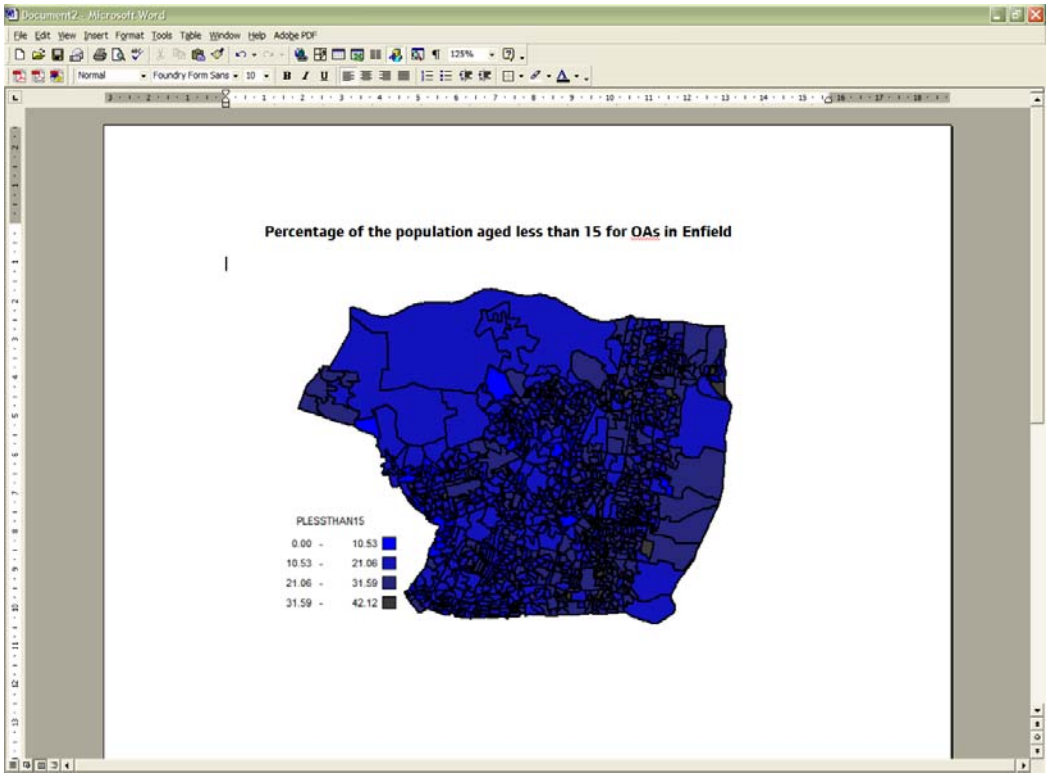
Command file:

```
input system file name = "C:\SASPAC\SYSFILES\/
2001 CENSUS AREA STATISTICS FOR OAS IN ENFIELD.SYS"
plessthan15 = (CS0010006+CS0010036+CS0010066) / CS0010001 *100.0
set countycode on
save ZONEID plessthan15
output CSV file with headers name = /
C:\SASPAC\INTERFAC\Module_5.CSV
end
finish
```

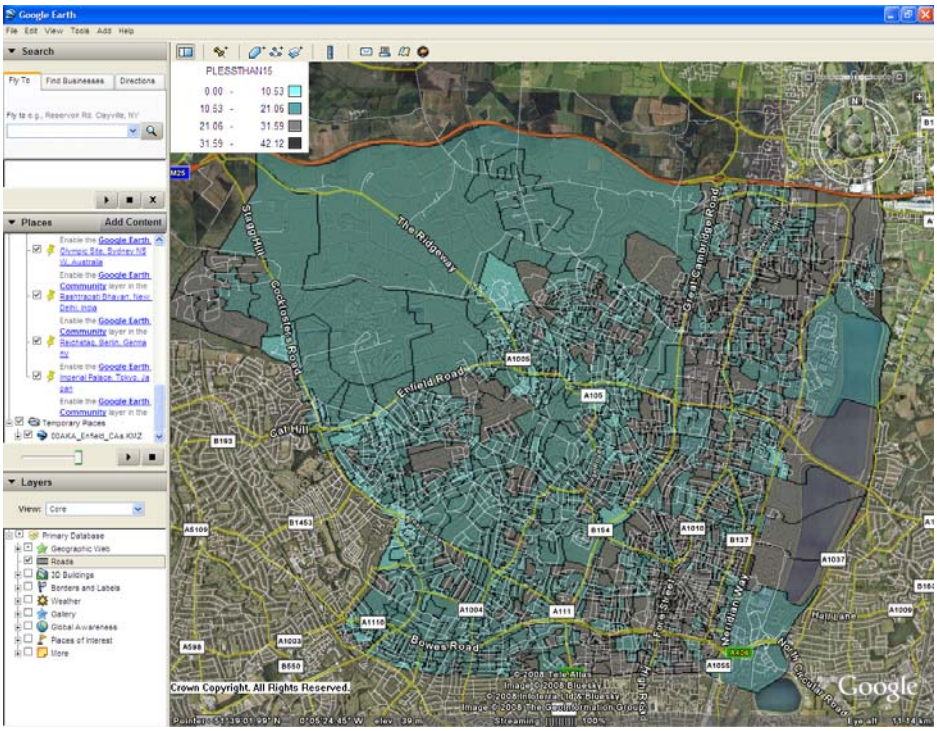
CSV file:

	A	B	C
1	ZONEID	PLESSTHAN15	
2	00AKGL0001	17.22	
3	00AKGL0002	21.61	
4	00AKGL0003	24.01	
5	00AKGL0004	9.32	
6	00AKGL0005	12.82	
7	00AKGL0006	22.18	
8	00AKGL0007	25	
9	00AKGL0008	10.98	
10	00AKGL0009	20.8	
11	00AKGL0010	19.22	
12	00AKGL0011	15.49	
13	00AKGL0012	16.81	
14	00AKGL0013	16.48	
15	00AKGL0014	15.24	
16	00AKGL0015	19.06	
17	00AKGL0016	16.54	
18	00AKGL0017	17.25	
19	00AKGL0018	16.62	
20	00AKGL0019	26.3	
21	00AKGL0020	24.49	

*Exported as a *.jpg to MS Word:*


Document2 - Microsoft Word
Percentage of the population aged less than 15 for OAs in Enfield
PLESSTHAN15
0.00 - 10.53
10.53 - 21.06
21.06 - 31.59
31.59 - 42.12

*Exported as a *.kmz to Google™ Earth:*



Module 6

Command file:

```
input system file name = "C:\SASPAC\SYSFILES\/
2001 CENSUS AREA STATISTICS FOR OAS IN ENFIELD.SYS"
perl15 = (CS0010006+ CS0010036+ CS0010066+/
CS0010101 ) / CS0010001*100.0
SELECT IF perl15>=35
print variables perl15
output print file = /
C:\SASPAC\REPORT\Module_6.PRN
end
finish
```

Print file: *(Page 1 of 1)*

```
 SASPAC CENSUS ANALYSIS PACKAGE

ZONE ID             PERL15
00AKGQ0003           35.39
00AKGQ0014           37.74
00AKGQ0024           37.79
00AKGR0033           36.36
00AKGS0015           36.20
00AKGS0024           37.77
00AKGY0002           35.16
00AKHC0014           45.05
00AKHF0017           36.06
```

Module 7a

Command file:

```
input system file name = "C:\SASPAC\SYSFILES\2001 CENSUS AREA/
STATISTICS FOR OAS IN ENFIELD.SYS"
input system file name = "C:\SASPAC\SYSFILES\2001 CENSUS AREA/
STATISTICS FOR WARDS IN ENFIELD.SYS"
read in series
new zone id = Zone001 name = New Zone 1
USING AREAS - oalevel 00AKGT0007 +(oalevel 00AKGW0001 to/
00AKGW0003) + oalevel 00AKGW0008*0.33 + wardlevel 00AKGT/
output system file name = C:\SASPAC\SYSFILES\Module_7a.SYS /
label=Training Module Exercise 7a
end
finish
```

Module 7b

Command file:

```
Input gazetteer file with labels name = C:\SASPAC\COMMAND\Enfield.gaz
existing zone cols 1 /
to 10 scale factor cols 21 to 24 new zone cols 13 to 18
input system file name = "C:\SASPAC\SYSFILES\2001 CENSUS AREA/
STATISTICS FOR WARDS IN ENFIELD.SYS"
input system file name = "C:\SASPAC\SYSFILES\2001 CENSUS AREA/
STATISTICS FOR OAS IN ENFIELD.SYS"
read in series
set zone echo on
output system file name = C:\SASPAC\SYSFILES\Module_7B.SYS/
label=Training Exercise Module 7B
end
finish
```

Module 8

Command file (Create FWK file):

```
input framework data file name = C:\SASPAC\FRWDATA\MYTABLE.fwd
output framework file name = C:\SASPAC\FRWORK\MYTABLE.FWK
end
finish
```

*Command file (Print new table)***:**

```
input system file name = "C:\SASPAC\SYSFILES\2001 STANDARD/
TABLES FOR DISTRICTS IN LONDON.SYS"
input framework file = C:\SASPAC\FRWORK\MYTABLE.FWK
include ualevel 00AC
print tables my01
output print file = C:\SASPAC\REPORT\Module8.PRN
end
finish
```

Print file:

Age Profile of Resident Population of:

Area Name: **Barnet**

Area ID: **00AC**

Total Population:	314,559
Population aged	
0 - 4	20,215
5 - 9	20,537

Module 9

Command file:

```
input system file name =
"C:\SASPAC\SYSFILES\KEY STATISTICS - OAS – CENTRAL
LONDON.SYS"
select if distmile(TQ32868024) le 1
* <<radial=1,532865,180245,I,mile,1,,>>
print variables with statistics  KS0010004 KS0150001 /    KS0150011
output print file = c:\saspac\report\Module9.prn
end
finish
```

Print file:

	KS0010004	KS0150001	KS0150011
TOTAL	90310	40360	9939
MEAN	282.22	126.13	31.06
STAND. DEV	107.04	40.71	19.97
MAXIMUM	1229	431	130
MINIMUM	100	63	5
MISS. VAL	0	0	0

Module 10

Step1 – create a new system file containing just the flows within the area of interest (Core Area)

Command file:

```
input system file name =
C:\SASPAC\SYSFILES\SWS2XX_WITH_LABELS.SYS
include 00AKGL to 00AKHH
output system file name =/
C:\SASPAC\SYSFILES\SWS2XX_Enfield_core.SYS label=SWS2XX /
London core
end
finish
```

Step 2 – rezone the areas outside the core area

Command file:

```
TEST 71
Input gazetteer file with labels name = /
C:\SASPAC\COMMAND\Ward_to_District_Enfield.gaz /
existing zone cols 1 to 6 new zone cols 15 to 18
input system file name = C:\SASPAC\SYSFILES\SWS2XX_Enfield_core.SYS
set zone echo on
output system file name =
C:\SASPAC\SYSFILES\SWS2XX_Enfield_Rezone.SYS label=SWS2XX /
London rezone
end
finish
```

Step 3 – Extract and map the number of persons working within the area

Command file:

```
input system file name = C:\SASPAC\SYSFILES\/
SWS2XX_Enfield_Rezone.SYS
include 00AKGL to 00AKHH
set in&out off
set countycode off
save ORIGID DESTID SWS2010001 SWS2010002 SWS2010003
output CSV file with headers name =
C:\SASPAC\INTERFAC\SWS2XX_Enfield.CSV
end
finish
```

CSV file:

	A	B	C	D	E	F
1	ORIGID	DESTID	SWS2010001	SWS2010002	SWS2010003	
2	00AKGL	00AKGL	812	545	267	
3	00AKGL	00AKGM	12	9	3	
4	00AKGL	00AKGN	12	3	9	
5	00AKGL	00AKGP	41	17	24	
6	00AKGL	00AKGQ	55	25	30	
7	00AKGL	00AKGR	42	26	16	
8	00AKGL	00AKGS	9	6	3	
9	00AKGL	00AKGT	33	15	18	
10	00AKGL	00AKGU	25	12	13	
11	00AKGL	00AKGW	51	10	41	
12	00AKGL	00AKGX	30	14	16	
13	00AKGL	00AKGY	3	3	0	

Data mapped in MapShore: